LAND ECONOMICS

a quarterly journal of
PLANNING, HOUSING & PUBLIC UTILITIES

CONTENTS — NOVEMBER 1956

Reports and Comments

Book Reviews

PUBLISHED QUARTERLY BY THE UNIVERSITY OF WISCONSIN
DURING THE MONTHS OF FEBRUARY, MAY, AUGUST, AND NOVEMBER

Publication Office: 121 South Pinckney Street, Madison, Wisconsin
Editorial Office: Sterling Hall, University of Wisconsin, Madison 6, Wisconsin
The contents of the *Journal* are indexed in the *Industrial Arts Index*
Information on *Preparation and Submission of Manuscripts* available on application to Editorial Office

Entered as second-class matter, January 3, 1938, at the post-office at Madison, Wis., under the Act of March 3, 1879. Acceptance for mailing at special rate of postage provided for in section 1103, Act of October 3, 1917, authorized October 12, 1922. Printed in the United States of America.

Subscription Rates: $6 a year; $2.00 a copy. (Single copies 1 year or more old; $2.25). Remittances may be made by personal checks, drafts, post-office or express money orders payable to *Land Economics*.

Agents of the *Journal* in Great Britain, B. F. Stevens & Brown Ltd., 28-30 Little Russell St. British Museum, London, W. C. 1.

LAND ECONOMICS

a quarterly journal of

PLANNING, HOUSING & PUBLIC UTILITIES

Published by The University of Wisconsin

LAND ECONOMICS

a quarterly journal of
PLANNING, HOUSING & PUBLIC UTILITIES

NOVEMBER
1956

VOLUME XXXII
NUMBER 4

Concepts Used as Economic Criteria for a System of Water Rights†

By S. V. CIRIACY-WANTRUP*

I. *Economic Criteria "In" and "For" Water Law*

ECONOMIC criteria are frequently implied in statutes concerned with water law, in judicial creation of water law through decisions in individual cases of controversy, and in administrative regulations by executive agencies—criteria "in" water law.[1] Well-known examples are concepts like "reasonable" and "beneficial" use, "waste," "surplus" of water, "maximum development," and "adequate compensation." A semantic analysis, from the economic point of view, of these and similar concepts would be interesting and useful. But an analysis of criteria "in" water law appears to cover only one aspect.

Economic criteria are common in semipopular and technical discussions, both in law and economics, when a system of water rights is considered as a whole—criteria "for" water law. In this country, there are three systems. First, in the eastern states, water rights are based on the riparian doctrine—with modifications in some states as, for example, North Carolina. Second, in the Great Basin and Mountain states, water rights are based on appropriation. Third, around the fringes of this heartland of the appropriation doctrine, the prevailing system of water rights is a "blend" exhibiting features of both doctrines—although in secular perspective the appropriation doctrine appears in the ascendency; this blend prevails in the Pacific Coast and the High Plains states.

In appraising these systems of water rights, a dichotomy of criteria is used. One criterion is exemplified by a set of concepts such as "security," "protection," and "rigidity" of water rights. The other criterion is represented by concepts like "flexibility," "adaptability," and "insecurity." Anyone familiar with the literature cannot fail to become impressed by the vagueness, plasticity, and contradiction which characterize the use of these concepts. An examination of their economic meaning is needed.

This dichotomy of criteria is applied jointly. In examining its application,

† Prepared for the *Symposium on the Law of Water Allocation in the Eastern United States*, Washington, D. C., October 4-6, 1956. The author wishes to acknowledge helpful comments by Wells A. Hutchins and Stephen C. Smith.

* Professor of Agricultural Economics, University of California.

[1] "Water law" will be interpreted here broadly to include contributions by the legislative, judicial, and executive branches of government.

one has to explore the gradations on the logical axis between the two poles and the resulting compromise in institutional arrangements. But also, one has to examine to what extent the two criteria can be applied together without such a compromise.

The consequences of institutional arrangements, if viewed over time, are complex. Applying a logical polarity to relations in reality does not always give a perfect "fit." Although it frequently happens—examples will be given below—that a change in institutional arrangements results in an increase in terms of one criterion and a decrease in terms of the other, this does not always happen. In other words, that the two criteria are poles logically does not necessarily mean that a change in institutional arrangements cannot be considered which results in increases in terms of one criterion without changes in terms of the other or that results in increases (decreases) in terms of both. In this, and in other aspects, our dichotomy is similar to that of "order" and "freedom" which has occupied students of jurisprudence for a long time and is not being neglected by economists of quite different "schools."[2]

The two criteria imply a problem area which is one of the most important and difficult, both for economic theory and policy. This is the problem area of "economic change" and of "dynamics" versus "statics" in economic discourse. This area is also the one in which the relations between law and economics raise some of the most acute and baffling issues.[3] Thus, focusing on these two criteria brings us to the core of the main theme.

Focusing on economic criteria used "for" water law does not mean that economic criteria used "in" water law are to be neglected. Both can be regarded as means of serving a common end—the "public interest." The public interest is the concept that connects criteria "for" with those "in" water law. Examining the economic meaning and implications of the former is not without relevance for the latter.

In examining concepts as criteria, the emphasis of this paper is on functional relations in economics, not on legal history or on normative meanings in law. The significance of the case law in this country is impressive—especially to one, as the present writer, who grew up in the legal climate of the Code Napoleon and its successors. But tracing historically the interpretation of economic concepts through the maze of case law is more a task for a student of law than for an economist. Likewise, it would be presumptuous for an economist to suggest how economic concepts should be interpreted in law.

A great deal is being written lately on the "integration" of law and economics. If by this term is meant that students in the two disciplines need greater understanding for each other's problems, tools, and limitations, one can wholeheartedly agree. If it is suggested by this term that concepts and processes of concept formation employed by economists should be transplanted to law (and vice versa), the prospective benefits would seem dubious. On the other hand, emphasis on the functional relations of concepts used as economic criteria "for" and "in" law may help in clarifying areas of common interest between two social science disciplines.

[2] John R. Commons, *Institutional Economics* (New York: The Macmillan Company, 1934); F. H. Knight, *The Economic Organization* (New York: A. M. Kelley, 1951), and *Freedom and Reform* (New York: Harper and Brothers, 1947). Lionel Robbins, "Freedom and Order," *Economics and Public Policy* (Washington, D. C.: Brookings Institution, 1955). (Brookings Lectures, 1954.)

[3] See: S. V. Ciriacy-Wantrup, "Some Economic Issues in Water Rights," *Journal of Farm Economics*, December 1955, pp. 875-885.

The "functional relations" to be studied in this paper may be indicated by two closely connected questions. First, one may ask, what are the economic implications—in the sense of logical and probable factual consequences—if concepts used as criteria "for" and "in" water law are interpreted and applied in certain ways. Second, one may ask, how far and why are these implications helpful or obstructive if certain economic objectives are sought. The economic interpretations and objectives selected should, of course, have relevance for actual problems of public water policy.

Although indicated by these questions, it may be well to point out explicitly at this time that the problem at hand will be viewed as one of positive rather than normative economics.[4] The consequences of this approach for the relations between economics and law will become apparent later (Section 6).

II. *Interpretation of "Security" of Water Rights*

To the economist, "security" of water rights means something different and much broader than their "protection" means to the student of law. The latter concept merely means protection against *unlawful* acts by others—as such acts are construed by the law. Such protection is always subject to the two major categories of "legal uncertainty," that is, to "rule uncertainty" and to "fact uncertainty."[5] Legal uncertainty, in this sense, is a characteristic of judicial decisions. Like other types of uncertainty, it also affects economic decisions.

Economists are inclined to disregard or underestimate the significance of legal uncertainty. For them, security of a water right connotes (1) protection against what I propose to call "physical uncertainty," that is, against variability over time of the quantity of water usable under the right due to seasonal or annual variability of "natural" run-off and ground-water recharge, and (2) protection against what I shall call "uncertainty of water tenure"[6] or, for short, "tenure uncertainty"; that is, protection against variability over time of the quantity of water usable under the right due to *lawful* acts of others—who may be individuals or groups, private or public.

We are concerned here with the relative degree of security resulting from different types of rights. Absolute security, that is, transformation of probabilities into single-valued expectations, cannot be obtained through water law. Furthermore, we are not concerned with the many other types of uncertainty which affect economic decisions—for example, uncertainties connected with variability over time of wants, technology, prices, and incomes. A few examples may illustrate differences between water rights if "security" is interpreted in the way just indicated.[7]

Under natural conditions a senior appropriative right is more secure than a junior appropriative right against physical uncertainty but not necessarily against tenure uncertainty. A water right in a high preference class is more secure than one in a low preference class against tenure uncertainty but not necessarily against physical uncertainty. A water right restricted by reservations in

[4] For a discussion of this differentiation see: Milton Friedman, *Essays in Positive Economics* (Chicago: University of Chicago Press, 1953).

[5] This terminology has been popularized by Frank. See: Jerome Frank, *The Law and the Modern Mind* (6th Ed.; New York: Coward McCann, 1948).

[6] Tenure uncertainty is not confined to water—and other "fugitive" resources—but is one of the most important economic forces affecting resource use. See: S. V. Ciriacy-Wantrup, "Capital Returns from Soil-Conservation Practices," *Journal of Farm Economics*, pt. 2, November 1947, pp. 1181-1196.

[7] For these examples, one assumes, of course, "other things being equal," that means in this case physical conditions and legal features other than those under consideration being the same.

favor of other users is less secure than these other rights against tenure uncertainty but not against physical uncertainty. An appropriative right is more secure against tenure uncertainty than a riparian right or a ground-water right under the correlative rights doctrine——assuming that only one system of rights applies to a water resource. Against physical uncertainty, however, an appropriative right is not necessarily more secure than a riparian or correlative right; this holds especially (but not solely) for a junior appropriative right.

In all these examples, the two main categories of legal uncertainty are also present. The degree of legal uncertainty may be quite different for different water right systems and individual water rights. For purposes of this paper, however, it is meaningful to proceed "as if" the degree of legal uncertainty were approximately the same.

If one speaks of "quantity of water usable" under a water right, one implies certain characteristics with respect to water quality. Over time, water quality is also subject to legal, physical, and tenure uncertainties. When water rights are defined in quantitative terms—as under appropriation and through adjudication under the riparian and correlative rights doctrines—reference should always be made to water quality.

Security of water rights in terms of water quality is no less significant economically than security in terms of water quantity. Frequently, the higher water uses—in terms of value product—require a higher quality of water than the lower uses. Furthermore, quality and quantity are directly interrelated: water of high quality—in terms of low mineral content, especially sodium and sulphates—can be reused, sometimes for more than one cycle. Such water can also be used to upgrade surface and ground water of lower quality which could otherwise be used only for lower uses or not at all. These interrelations of quality and quantity are especially significant when water must be transported over great distances. Such transport may be economically feasible only for high quality water. It follows, from the interrelations of quantity and quality just explained, that in setting minimum standards for water quality, problems of reuse of the same water and upgrading of other waters need to be considered. Minimum standards based on suitability for one cycle of direct ("unmixed") use alone may be, economically speaking, too lenient.

Sometimes, the only security of water rights in terms of quality is their protection against unlawful acts of water pollution (including, for purposes of this paper, contamination and nuisance). In many states, antipollution laws are being strengthened. More federal action is also being considered.[8] Dealing specifically with these attempts would require a separate paper.

Physical uncertainty can be significantly reduced through physical means, that is, mainly through storage of water—above *and* below ground—for season to season, from year to year, and over a period of years. A reduction of physical uncertainty in this sense produces benefits for water users and others. Whether these benefits are "net" depends on the costs of constructing, maintaining, and managing storage.

A system of water rights may impede construction of storage facilities by limiting development to a vaguely defined "safe yield" of a water resource and by making cyclical management of storage

[8] The federal government is no newcomer in this field. See, for example, the legislation known as the Caminetti Act, 27 *U. S. Stat. at L.* (1893), 507, as amended by 30 *U. S. Stat. at L.* (1898), 631; 30 *U. S. Stat. at L.* (1899), 1148; 31 *U. S. Stat. at L.* (1900), 631; 34 *U. S. Stat. at L.* (1907, 1001; 48 *U. S. Stat. at L.* (1934), 1118; 52 *U. S. Stat. at L.* (1938), 1040.

capacity and coordination of surface and subsurface storage more difficult.

The appropriation doctrine can easily be used to limit water development. In several western states this doctrine has actually been used with this objective and result in the case of ground water. With respect to surface water the actual facts are somewhat different. Most western surface waters are "over appropriated." Sometimes, appropriation covers a multiple of the average flow during the season of use.

After storage capacity has been provided and is managed with a view to reducing physical uncertainty, the relative economic status of appropriative rights changes without changes in their relative legal status. Priority in time, in conjunction with the quantitative definition of appropriative rights, limits the number of rights that can be served with the regulated flow. Rights exceeding this flow become, economically speaking, less meaningful the better the flow is regulated. For rights that can be served with the regulated flow, the new situation, in terms of economics, is not greatly different from that prevailing under a water delivery contract where a limited number of users are equal in right—although the quantities to which their contract entitles them may differ.

Thus, the differentiation of water rights, so characteristic for the appropriation doctrine, has undergone a shift in its economic implications: within limits a junior right is no longer less secure against physical uncertainty than a senior right. Increasingly, storage is provided by large public or semipublic projects and managed by government agencies—public districts, state water departments and boards, and federal bureaus.

In contrast to appropriation, riparian rights are coequal in law.[9] An economic implication of this equality is that their legal status is not a cause for differentiation with respect to their security against physical uncertainty. This holds both for the riparian right to surface waters, as it is generally applied in the eastern states, and for the application of the riparian right to ground water through the correlative rights doctrine as, for example, in California.

This lack of differentiation on the basis of legal status does not necessarily mean that there are no differences between individual riparian or correlative rights in security against physical uncertainty. Applying legal equality through adjudication to a given stream system or ground-water basin is a time-consuming and costly process. For a while, at least, individual riparian and correlative rights may differ in security against physical uncertainty. Such differences are caused not by legal but physical facts such as relative location of diversions along a surface stream and the location of wells with respect to the source of recharge and the geologic characteristics of a ground-water basin.

Differences between individual water rights (and between water-right systems where they exist side by side in the same general area) in security against physical uncertainty may obstruct or retard construction and coordinated cyclical management of storage. Individuals and groups who "have" greater security may be reluctant to join those who "have not" if a portion of the costs of reducing physical uncertainty is incident on the former.

[9] They are coequal within the two traditional preference classes of "natural" and "artificial" uses—except that an upstream user need not share with downstream users if the flow of the stream is only sufficient to satisfy his own natural uses.

There are a number of institutional approaches to this problem. One approach is based on a type of public district through which water rights are acquired, or which acts as agent in utilizing them, for coordinated cyclical management of surface and ground water. This approach is now being investigated in ground-water studies[10] for Santa Clara County, California. In many parts of the West water demand is increasing so rapidly that the group of "haves" (in terms of security of water rights against physical uncertainty) is becoming smaller absolutely and, of course, even more so relative to the number of "have nots." The formation of public districts and of other institutional arrangements for coordinated water management is thereby facilitated.

Turning now to security of water rights against tenure uncertainty, it was mentioned earlier that appropriated rights are more secure in this respect than riparian rights or ground-water rights under the correlative rights doctrine. The former are clearly defined in priority, quantity, period of use, points of diversion, and in other ways. The latter are coequal, and quantitative definition depends on adjudication which is in terms of shares and subject to the restriction that there shall be reapportionment if the conditions upon which the original apportionment was made change sufficiently to justify it.

There are, however, several factors which modify such a general comparison of water-right systems. The first factor is prescription which operates under all water-right systems—with some differences in economic significance (Section IV). The second factor is differentiation of preference classes based on purpose (kind, type) of water use. This factor also operates under all water-right systems but under appropriation only *before* rights are vested and under emergencies. Customary preference classes are "natural" and "artificial" under the riparian doctrine and "domestic," "municipal," "industrial," "agricultural," and "recreational" under appropriation. The third factor is a restriction ("reservation") on water rights to yield to future water demands by others. This factor operates only under appropriation and affects future (not already vested) rights. Water reservations are frequently used in favor of certain preference classes. But the connection with preference classes is not a necessary one. In California water reservations in favor of whole regions are in force regardless of preference class.

There can be little argument on economic or any other grounds that domestic and municipal uses deserve special consideration as far as security against tenure uncertainty is concerned. This type of use can frequently prevail over other uses by eminent domain proceedings. Still, the highest preference ranking and water reservations in favor of this use are desirable because other uses, for example agricultural, may also be organized as public districts. By adequate requirements for the treatment of sewage, domestic and municipal use can be made largely nonconsumptive. Agricultural use, on the other hand, is largely consumptive.

[10] Stephen C. Smith, "Problems in Using the Public District for Ground-Water Management," *Land Economics*, August 1956. For other studies in this series, see: Patricia McBride Bartz, with a foreword by S. V. Ciriacy-Wantrup, *Ground Water in California, The Present State of Our Knowledge* (Berkeley: University of California, College of Agriculture, Agricultural Experiment Station, September 1949), 67p. (Giannini Foundation Ground Water Studies No. 1.) Processed; Herbert J. Snyder, with a foreword by S. V. Ciriacy-Wantrup, *Ground Water in California, The Experience of Antelope Valley* (Berkeley: University of California, Division of Agricultural Sciences, Agricultural Experiment Station, February 1955), 171p. (Giannini Foundation Ground Water Studies No. 2.) Processed.

An economic argument can be made in favor of abolishing the usually[11] lower preference rating of industrial uses relative to agricultural. The average[12] value product of consumptive use is higher in industry than in agriculture. Furthermore, in the western states total consumptive use in industry—although increasing—is still relatively small as compared with that in agriculture.

In ascertaining the value product of recreational uses—which are largely nonconsumptive—one faces the difficult problem of extra market values. There are cases, however, in which the economic argument suggests a change in the present preference ranking of recreational uses. Generally, this ranking is the lowest. Under some statutes, recreational uses are not even recognized as "reasonable and beneficial."

In spite of the existence of a favorable economic argument for abolishing the usual statutory ranking of industrial and agricultural uses and for other changes in preference classifications, inferences with respect to public policy would be premature without considering the criterion of flexibility jointly with that of security. The economic implications of water reservations likewise cannot be fully appraised without considering the criterion of flexibility. An interpretation of this criterion will be undertaken in a later section (Section IV). Before this can be done, our interpretation of security of water rights must be related to "protection of investment" in water resources development.

[11] Significant exceptions occur, for example, in Texas.

[12] We are using the word "average" because each of the two preference classes covers water development by different users for different products in different time intervals. In the present context, we refer, of course, to future "additional" water development. A change in preference classification does not affect already vested appropriative rights. Under the riparian doctrine, industrial and agricultural uses are both "artificial" and, therefore, coequal.

III. *Security of Water Rights and Protection of Investment*

The relations between security of water rights and investment in water resources development are generally the main point of emphasis when the economic implications of differences in individual water rights and whole water-right systems are discussed. This emphasis is justified. Most economic implications of security which were mentioned in the preceding section are in this area. The objective of the present section is to state some of these relations more explicitly and to consider them in the light of the concept, "protection of investment."

In economic theory, "investment" and its corollary "disinvestment" refer to value changes in total capital of individuals or whole social groups as a result of differences between income and consumption. This is not what is meant here. In the present context, "investment" refers to what the economist would call the value of particular durable physical "assets."[13] The value of durable physical assets depends on the flow of net income which the assets are expected to "yield" over time.[14] Assets themselves, however, refer to the present; and the income flow which determines their value is subject to a time discount and an allowance for uncertainty. Thus, we are concerned here with the protection against physical and tenure uncertainties to which this income flow is subject. The degree of such protection differs greatly for water rights with different security—as explained in the preceding section.

[13] One may differentiate between three forms of assets: (1) physical assets (natural resources, improvements, equipment, inventories); (2) money assets (securities, loans, cash); and (3) personal assets (labor, skills, and "good will" of the individuals or groups who hold assets).

[14] The resale or "scrap" value of durable physical assets at the time they are sold or scrapped may be formally included in this flow.

At first glance, it might be expected that, other things being equal, a greater or smaller security of water rights will result in an increase or decrease of investment in water resources development. This inference needs some scrutiny.

Frequently a greater degree of security for some water rights necessarily entails a smaller degree of security for other rights. For example, dividing appropriative rights on a given surface stream into senior and junior rights increases security against physical uncertainty for the former but decreases it for the latter. If a certain type of use—such as municipal—receives a higher preference ranking than another—such as agricultural—or if the rights of the latter are restricted by water reservations in favor of the former, security against tenure uncertainty is increased for water rights held by the municipality but decreased for rights held by the agricultural users. This situation does not make the above inference invalid in terms of investment by individuals. But it should not be applied to aggregate investment in water resources development of a given surface stream, ground-water basin, or region. From the standpoint of public water policy, such aggregates of investment are important.

The statement was made in the preceding section that generally (that is, without taking account of modification through prescription, preferences, reservations, and physical factors) appropriative rights give greater security against tenure uncertainty than riparian and correlative rights. This comparison refers to all water rights on a surface stream, in a ground-water basin, or in a region. In this case, therefore, the inference with respect to investment is valid for aggregates of investment. For this reason the conclusion that the appropriation doctrine favors investment in water resources development—if compared with the riparian and correlative rights doctrines—is justified. One should keep in mind, however, that only protection against tenure uncertainty, in the defined sense is involved.

When one compares different water rights and water-right systems in terms of implications for investment, one touches on aspects of the concept "adequate compensation." This occurs, for example, if the degree of protection against tenure uncertainty is affected by prescription, preferences, and reservations. Prescription does not involve compensation according to law. Preferences and reservations involve compensation under some laws but not under others.[15] Regardless of the legal aspects, one may raise the question whether and under what conditions compensation might be considered in public policy as a problem of economics.

Let us assume that a reservation exists on the flow of a surface stream and that a municipality holds the reservation and will not need the water for 20 years. During this period the water is available for temporary appropriation by other users. Let us assume that the only alternative use is agricultural. Let us assume further that such use involves considerable expenditure for diversion and storage dams, main canals, a distribution system, land leveling, and other durable

[15] The constitutions of Idaho and Nebraska grant preferences in time of scarcity of water, first to domestic uses, second to agriculture, but make exercise of the right contingent on payment of compensation. The Colorado constitution grants similar preferences without mention of compensation but the Colorado Supreme Court has held that, despite that provision, full compensation is required. Statutes of Oregon and Utah give similar preferences in time of scarcity without mention of compensation. In Texas, reservation does not involve compensation, although there is a difference of opinion as to the validity of the statute granting blanket reservations to municipalities. In California, a municipality holding a reservation must compensate the temporary appropriator. For these and other differences in state laws, see Wells A. Hutchins, *Selected Problems in the Law of Water Rights in the West* (Washington, D. C.: 1942), especially pp. 337-358 (U. S. Department of Agriculture Miscellaneous Publication No. 418.)

improvements. A private user will make these expenditures only if they seem warranted by the income stream that the durable assets are expected to yield. From the standpoint of the private user, the duration of the income stream is uncertain because of his water tenure. Under these conditions, the expenditure may not be forthcoming, and the water may go unutilized for 20 years.

In such a situation a guarantee of compensation for nonrecovery (because of termination of water tenure) of expenditures plus a sufficient profit margin would offset the deterrent to a private user to develop the available water on a temporary basis. Whether "protection of investment" in this sense would be economically warranted from the standpoint of public water policy is not self-evident but can be ascertained by benefit-cost analysis.[16]

In such an analysis some benefits are considered which the private user must leave out of account. Such benefits may, for example, result from flood control or ground-water recharge. Benefits yielded after 20 years are also considered. For example, dams and other facilities may be usable by the municipality although may not be legally required to pay for them. Further, some costs of construction which the private user must consider appear smaller in benefit-cost analysis—for example, labor costs of construction in a period of unemployment under sticky wage rates.

The foregoing argument in favor of "protection of investment" in water resources development is based on two necessary conditions: (1) that expenditures for durable assets are in the public interest although they may not be economical for private water users, and (2) that the most economical alternative for public policy to develop water resources is a guarantee just sufficient to induce private development. There is no implication in this argument that "protection of investment" per se is in the public interest.

[16] S. V. Ciriacy-Wantrup, "Benefit-Cost Analysis and Public Resource Development," *Journal of Farm Economics*, November 1955, pp. 676-689.

IV. *Interpretation of "Flexibility" of Water Rights*

In interpreting flexibility of water rights one thinks first of all of "legal flexibility." This is a corollary to "legal uncertainty," the two main categories of which were mentioned above (Section II).

Legal uncertainty may be regarded as the price that must be paid for obtaining legal flexibility. There has been considerable discussion within the legal profession on whether or not the product—a law responsive to the needs of the community—stands in fair relation to its price. During the twentieth century the trend in legal thinking has been to answer this question in the affirmative. To be sure, there are differences among prominent legal thinkers with respect to the desirable degree of flexibility in particular areas of law—such as property rights; but all agree (including non-American jurists) that flexibility is necessary and that the American development of the Anglo-Saxon common law is better suited than other legal systems for approaching the most desirable degree of flexibility through trial and error and step by step. Within jurisprudence this degree at a given time and place will remain the subject of perpetual discussion, reappraisal, shifting of emphasis, and ambivalent attitudes. Far from being a weakness, such a condition would appear a normal or even necessary one for an effective contribution by the law to "social engineering"—to use a term

popular with Roscoe Pound[17] and other students of law.

Within economics we are concerned not so much with legal flexibility as with the needs themselves to which the law responds. In economics, therefore, the interpretation of "flexibility of water rights" differs from that of "flexibility of water law" as discussed in jurisprudence, and an increase in flexibility of water rights is not necessarily incompatible with an increase in their security—as stated in the beginning (Section I). The needs of particular interest in connection with flexibility of water rights are created by economic change.

The impact of economic change upon water use may be divided conceptually into two parts— which in reality, and for water policy, are not independent: (1) a change in aggregate development of water resources within a region (country, state, watershed, ground-water basin), and (2) a change in allocation of water resources between regions, uses (municipal, agricultural, industrial, recreational), and users (individuals, firms, public districts, government agencies). The criterion "security of water rights" is significant for the former change—as discussed in the preceding two sections. The criterion, "flexibility of water rights," as interpreted here, focuses on those aspects of water rights which facilitate or obstruct changes over time in the allocation of water resources between regions, uses, and users. Several such aspects must be considered.

There is first the "transferability" of water rights. Water rights, like other property rights, can be transferred in various ways. Voluntary transfer of water rights through buying and selling —with land if they are "appurtenant" or without land if they are not—is not uncommon. Appropriative rights are better suited for such transfer than riparian rights because the former are clearly defined in quantity, priority, points of diversion, and other ways. Transfer of riparian and correlative rights generally requires also transfer of land.[18] Furthermore, they are not clearly defined quantitatively and insecure against physical and tenure uncertainties (Section II). With respect to transferability, therefore, the appropriation doctrine favors flexibility of water rights in the course of economic change.

The market for water rights—and especially its "imperfections"—is an interesting topic for the economist. This paper, however, is more concerned with flexibility of water rights as a whole. Transfer of water rights through the market mechanism is only one aspect. Involuntary transfer of water rights is probably more important in number of transactions and quantity of water involved than transfers through the market.

Involuntary transfer may, first of all, be accomplished through condemnation for public use. This process, by constitutional provision, involves compensation. Appropriative rights are better suited for this process than riparian and correlative rights for the same reasons as were mentioned in connection with voluntary transfer. Furthermore, it is usually sufficient to condemn only a few appropriative rights with high priority. Under the riparian doctrine it is frequently necessary to condemn *all* riparian rights. In this comparison, we assume of course that only one system of rights applies to a water resource (Section II).

Second, involuntary transfer can be accomplished through prescription. This

[17] Roscoe Pound, *The Spirit of the Common Law* (Francestown, New Hampshire: Marshall Jones Company, 1921).

[18] Riparian rights may be waived by a grant to nonriparians. Such grant is not effective against other riparians and is not really a transfer of the riparian right. But from the economic standpoint it is just as effective. It frequently happens that riparian rights are bought up or condemned.

process of transfer does not involve compensation. As we know (Section II), prescription operates under all water-right systems; but there are differences in its economic significance under various systems.[19] Prescription takes time—from 3 to 20 years (5 years in California), depending on the statute of limitation—to ripen into a vested prescriptive right. Adverse and open beneficial use, the condition for prescription, is more likely to persist unobjected to for such a period against downstream riparian rights and correlative rights than against appropriative rights. A prescriptive right can in turn be lost through prescription by others. Prescription, therefore, is not irreversible but a potentially always present and economically interesting process of transfer.

Abandonment and forfeiture of water rights—operating only under appropriation—may also be regarded as processes of transfer. Their present over-all significance for flexibility of water rights is minor. However, through defining conditions of forfeiture—for example, through legislative, judicial, and executive defining and redefining of "beneficial' use—the economic significance of this process could be increased.

Voluntary and involuntary transfer of water rights is only one aspect of flexibility. There is fairly general agreement that, aside from the aspect of transferability, a water right system based on the riparian and correlative rights doctrines if compared with a system based on appropriation, contains elements of greater flexibility. Under the riparian doctrine, new uses created by economic change compete on an equal legal basis with older uses and obtain rights that are no less secure against physical uncertainty than older rights. This is a corollary to the tenure uncertainty so characteristic for riparian and correlative rights (Section II). Under the appropriation doctrine, on the other hand, new uses can obtain only inferior rights in terms of security against physical uncertainty. As just noted, however, the economic significance of such flexibility, inherent in riparian and correlative rights, is reduced over time for a given region through prescription.

Preferences and reservations are important for flexibility because of their obstructive influence. They tend to continue into the future the favorable treatment of certain uses or users on the basis of present economic conditions. Preferences and reservations do not obstruct growth itself. Quite the contrary, they are designed to facilitate growth. They facilitate, however, only the growth of that use which is deemed to deserve preferential treatment on the basis of present economic conditions. They obstruct the growth of other uses. Economic change may well require a change in the ranking of uses. As already suggested (Section II), preferential treatment is unobjectionable for domestic and municipal uses. On the other hand, extension of preferences to other uses—for example, in favor of agricultural against industrial—or extension of reservations to all users of one region against users of another—for example, in favor of counties and watersheds of "origin" against those of "destination"[20] obstructs changes

[19] The differences suggested here do not involve the legal problem of whether or not a prescriptive title "good against the world" can be established without the necessity of valid statutory appropriation. For different conclusions on this point, see: Russell R. Kletzing, "Prescriptive Water Rights in California: Is Application a Prerequisite?" *California Law Review*, September 1951, pp. 369-376; Delger Trowbridge, "Prescriptive Water Rights in California: An Addendum," *California Law Review*, December 1951, pp. 525-527; Gavin M. Craig, "Prescriptive Water Rights in California and the Necessity for a Valid Statutory Appropriation," *California Law Review*, May 1954, pp. 219-242.

[20] State of California, *Water Code*, 1953 (Sacramento: California Printing Division), Section 10505, p. 195; *Ibid.*, Section 11460, p. 205; See also: Ciriacy-Wantrup, "Some Economic Issues in Water Rights," *op. cit.*

over time in the allocation of water resources.

The aspects of flexibility discussed so far relate to statutory water law. It would be a serious mistake to overlook the great significance of the case law for flexibility. As already stated (Section I), concepts are used as criteria "in" water law that imply an economic appraisal. These concepts are rather plastic and give considerable scope to the judicial development of water law through decisions in individual cases of controversy. Concepts like "reasonable" and "beneficial" use, "waste," "surplus" of water, "maximum utilization" are interpreted and reinterpreted by the courts continuously in the light of changing economic conditions. Over time, the result has frequently been a change in the allocation of water resources between uses and, through it, between users.

Besides taking economic change into account in deciding controversies, the courts are directly concerned with the transfer of water rights through eminent domain. Condemnation of water rights for public use is well established in all states. Going much farther than that, the State of Washington permits any person to condemn a lower water use for a higher use.[21] For this purpose any beneficial use is declared to be a public use. The courts determine which use is the higher one.

Under the American form of government, the contribution of flexibility by the executive branch of government is probably smaller than that of the legislative and the judiciary because of the constitutional issue of "delegation of power." In some states, however, the executive agency charged with the administration of appropriation statutes is given considerable discretion in granting and conditioning appropriation permits. Such discretion has been upheld by the courts.[22]

This section has considered the economic meaning of flexibility of water rights and the institutional possibilities through which flexibility may become operative in reality. There remains the task of considering flexibility jointly with security as economic criteria from the standpoint of public water policy.

V. *Welfare Economics and Water Allocation*

In an attempt to develop criteria for public policy, economists have developed a branch of normative economics called "welfare economics." More recently this branch has become known as the "new" welfare economics to emphasize its development in England and the United States since the 1930's. Its essential problems were recognized and its relevant theorems developed in the 1890's by Pareto.[23]

In formulating policy criteria, welfare economics takes explicit account of differences in individual preferences and

[21] Rev. Code Wash., sec. 90.04.030.

[22] Sections 1253 and 1254 of the California Water Code provide that the Department of Public Works shall allow appropriations under terms and conditions which "in its judgment" will best serve the public interest in water conservation. However, in acting upon applications the Department shall be guided by the policy that domestic use of water is highest and irrigation next highest. The California Supreme Court, in *East Bay Municipal Utility Dist.*, v., *State Department of Public Works*, 1 Calif. (2d) 476, 479-481, 35 Pac. (2d) 1027 (1934), upheld the action of the state agency in inserting in a permit, pursuant to these statutory provisions, the following condition: "The right to store and use water for power purposes under this permit shall not interfere with future appropriations of said water for agricultural or municipal purposes." In a more recent decision, in *Temescal Water Co.* v. *State Department of Public Works*, 44 Calif. (2d) 90, 99-101, 280 Pac. (2d) 1 (1955), the court held that the cumulative effect of statutory changes had been to create a type of proceeding greatly different from that considered in some earlier decisions. In carrying out its present duty, held the court, the Department of Public Works exercises a broad discretion in determining whether the issuance of a permit will best serve the public interest. That determination requires an administrative adjudication. If issuance of the permit is protested as the statute authorizes, the administrative decision may be made only after a hearing of the protest. The decision is subject to judicial review by way of writ of mandate. These two decisions were called to the author's attention by Wells A. Hutchins.

[23] Vilfredo Pareto, *Cours d'Economique Politique* (Lausanne: F. Rouge, Librarire-Editeur, 1897).

incomes and of the resulting problems in aggregating individual utilities. It is an economic axiom that the marginal utility of individual income decreases with increasing income. There is no agreement among economists on whether and in what sense—ordinally or cardinally—individual utilities can be compared; but welfare criteria that avoid interpersonal comparisons are generally preferred.

Classical and neoclassical economists were well aware of these problems.[24] They, however, focused on an increase of real aggregate national income as the main criterion of economic welfare.[25] Pareto's views were not in conflict with this emphasis because he believed—supported by historical experience as he saw it—that an increase of national income and greater equality of income distribution tended to be associated. In this case an increase of national income means also an increase of economic welfare according to Pareto's criterion at least under some generally accepted assumptions.

The positive correlation between changes of national income and of equality of income distribution—sometimes called "Pareto's law"—was challenged by Pigou[26] and others; but Pareto's welfare criterion is independent of his "law." This does not imply that the correlation noted by Pareto does not exist nor that Pareto's criterion is of greater significance for economic theory than his "law." Quite the contrary, one may wonder whether the great intellectual effort of the last 20 years which has been invested in developing Pareto's criterion might not have yielded greater dividends, in terms of knowledge as well as welfare, if it had been employed for further investigation of Pareto's law and of the problems associated with the increase of national income.

The Pareto criterion says that a change that makes at least one individual better off and leaves no individual worse off represents an increase of welfare. This criterion is usually interpreted to mean that welfare is increased by a change rendering it "possible" to make at least one individual better off and leave no individual worse off by compensating the losers. Most of the discussion in the new welfare economics deals with this compensation principle.

The Pareto criterion "without" compensation is so restrictive that it has little relevance for an appraisal of public policies—even if it could be practically applied. There are scarcely any policies which make nobody worse off. Furthermore, if there were such policies, the criterion would be ineffective for choosing between more than one alternative to the status quo. The Pareto criterion "with" compensation is not so restrictive but its application is even less practical.

The Pareto "with" criterion is conceptually not identical with the criterion "increase of national income." But the latter criterion may be regarded as a practical, first approximation to the former, provided that the policy under consideration does not appreciably increase inequality of income distribution; and provided further that there are other policies in operation which work independently and continually in in the direction of greater equality of income distribution. Such policies are, for example, progression in income and property taxes, high inheritance taxes, and "social welfare" legislation in the

[24] The first edition of Alfred Marshall's *Principles of Economics* appeared in 1890, seven years before publication of Pareto's main work in French. Marshall mentions Pareto only in passing and in a different connection.

[25] When comparing national income at different points of time and for different countries, per capita figures are used. In appraising alternative policies, it is more useful to focus on aggregate income.

[26] A. C. Pigou, *The Economics of Welfare* (London: Macmillan and Company, Limited, 1938).

narrower sense (relating to old age, invalidity, unemployment, minimum wages, and public health, education, and so on). In some practically important cases these two conditions can be regarded as fulfilled when considering resource policies in modern western societies.

Accepting an increase of national income as an economic criterion for public water policy does not imply that application of this criterion faces no theoretical and practical difficulties or that it is the most useful criterion under conditions where economic change and uncertainty are the central problems. We shall return to these problems in the concluding section (Section VI).

The contribution of welfare economics has been a clarification of the theoretical meaning (or absence of it) of a social welfare function and social indifference curves and of the difficulties (or impossibility) of applying the Pareto criterion in actuality.[27] The disservice of welfare economics has been that its terminology is used by economists and others without pointing out these theoretical and practical difficulties. The false impression is created that a simple criterion is available that can be used for legislation, court decisions, administrative regulation, and social planning in general. In the field of water allocation policy, such use of welfare economics can best be shown by an example.

In a recent paper, optimum water allocation in social planning is analyzed by superimposing smoothly convex social satisfaction indifference curves on a single production possibilities curve equidistant from the point of origin.[28] Anyone with a little knowledge of high school geometry has no difficulty in locating accurately a point of maximum social satisfaction in water allocation between two users.

No information is given as to how the social satisfaction indifference curves can be determined theoretically and computed in actuality. A higher indifference curve does not become "Pareto-better" by word magic. The suggestion that such curves could be used for water allocation in legislating and planning must be regarded as not warranted by the state of welfare economics.

For some time, cost and revenue indifference systems have been in use to analyze decision making in firm economics.[29] But applying a single equidistant possibilities curve to policy decisions raises questions no less serious than the use of social satisfaction indifference curves.

The ancestry of a single production possibilities curve, the apparent simplicity of which has made it rather popular recently, can be traced to two basic assumptions of programming.[30] These assumptions are, first, existence of limitational factors—especially capital—and, second, independence of decisions regarding the intensity of each process and decisions regarding the combination of processes. No information is given as to what limitational factors are assumed. From the standpoint of policy, capital and other factors are limitational only under narrowly defined short-run static assumptions. Under such assumptions, water allocation through legislation,

[27] For the last (but probably not final) word in this clarification, see: Paul A. Samuelson, "Social Indifference Curves," *Quarterly Journal of Economics*, February 1956, pp. 1-22. This article cites the significant previous literature.

[28] Earl O. Heady and John F. Timmons, "Economic Framework for Planning and Legislating Efficient Use of Water Resources," presented at the Seminar on Iowa's Water Resources, sponsored by the Agricultural Law Center, College of Law, State University of Iowa, and the Division of Agriculture, Iowa State College, Ames, Iowa, March 26, 1956, 22p. Processed.

[29] S. V. Ciriacy-Wantrup, "Economics of Joint Costs in Agriculture," *Journal of Farm Economics*, December 1941.

[30] The reader with no technical training in economics may want to consult an easily understandable explanation of the assumptions and techniques of programming, for example: James N. Boles, "Linear Programming and Farm Management Analysis," *Journal of Farm Economics*, February 1955.

court decisions, and administrative regulation has little meaning.

Under long-run static and under dynamic assumptions, policy decisions regarding water allocation are not independent of policy decisions regarding water development (Section IV). A single production possibilities curve "assumes away" the essential problem of the meaning and the determination of marginal costs of water development. Water allocation policy deals with a whole system of cost indifference curves. Which one is relevant can be ascertained only *after* comparing them with a system of "revenue" indifference curves. The optimum point of water allocation then becomes a curve of "optimum direction" using the terminology of vector analysis.[31] This curve is not monotonic and under assumptions approaching reality it is a space curve.

The equidistant feature of the single possibilities curve implies that changes in the water use vector have no influence upon production possibilities. By this implication one "assumes away" another essential problem of water allocation. This problem is created by great differences in water quality requirements between water uses. This problem presents itself in economic terms through important relations between quality, quantity, and costs. Some of these relations were indicated above (Section II).

One must conclude, therefore, that a single equidistant production possibilities curve has no meaning for water allocation policy—whatever its use may be in firm economics.

The foregoing example is of some interest because it combines the approaches and techniques of the two most important branches of normative economics—welfare economics and firm economics—and attempts to derive from such a combination criteria for public water policy. The questions raised so far with respect to this attempt are overshadowed in their implications for water policy by another: Is it conceptually useful to make the maximization principle the basis of economic criteria for pursuing the public interest?

VI. *Economic Criteria and the Public Interest*

The maximization principle is applied in normative economics; first, as efficiency criterion for limited operations under restrictive assumptions and second, as the assumed over-all objective of individuals and groups.

As efficiency criterion, the maximization principle is used, for example, in finding the optimum output under given cost and revenue functions and also in determining minimum costs for each output under given production functions and given price schedules of productive factors, that is, in determining a cost function. For these and similar purposes the maximization principle is necessary. There can be no disagreement on the usefulness of such operations. One may recall this application of the maximization principle "efficiency economics" or, more appropriately in some cases, "efficiency engineering."

If applied as the assumed over-all objective of individuals and groups, on the other hand, the maximization principle is a construct—a scientific fiction.[32] It is useful in economics, especially in modern western culture, if employed in connection with another construct—the "firm." Frequent references in recent economic literature to maxima of individual and social satisfaction indicate

[31] S. V. Ciriacy-Wantrup, "Economics of Joint Costs in Agriculture," *op. cit.*, p. 794.

[32] Next to mathematics and law, economics is the discipline in which scientific frictions are most common. But the natural sciences, especially modern physics, frequently employ fictions.

that the maximization principle is more and more applied as a fiction.

A fiction is permissible in science if its character is clearly understood. A fiction is a deliberate, conscious deviation from reality. A fiction, however, is not a hypothesis or theory. By itself a fiction is not intended to be validated by testing with empirical evidence. But a scientific fiction should be useful as a stimulus for or as a part of hypotheses and theories which can be tested. That means the test of a scientific fiction is its conceptual usefulness, its expediency in understanding, explaining, and predicting reality. A fiction becomes mere dogma and therefore unscientific if its two characteristics —consciousness of its fictional nature and conceptual usefulness—are obliterated. There are many examples in the history of science of fictions changing into dogma.

One may wonder whether or not the maximization principle has sometimes become dogma in economics. There is increasing emphasis on techniques which facilitate greater numerical accuracy in the determination of optima for the firm. These same techniques are then used for maximizing social satisfaction of whole groups with no conceptual gain and at the expense of "assuming away" essential economic relations; an example of such use was just given (Section V).

It was suggested above that, under certain conditions, an increase of national income may be accepted as a criterion for resource policies. The Pareto criterion, likewise, is suited only for ascertaining whether or not an increase of social welfare has occurred but not for determining a maximum.

The criterion "increase of national income" can be employed effectively in appraising water policies of more limited scope; for example, in appraising an individual water development project or in deciding a particular case of controversy in water allocation. This is the approach of benefit-cost analysis referred to earlier (Sections II and III). Limitations on the applicability of this criterion are imposed by a number of theoretical and practical difficulties some of which can be overcome only by restrictive assumptions.

The quantities of goods and services making up the national income must be evaluated (weighted) in order to be aggregated. The weights used—market prices and unit values derived indirectly from prices and in other ways—are affected by income distribution and by the host of institutions which influence this distribution. Both value weights and quantities are affected by market form. Policies to be appraised may change income distribution and market form. Such an appraisal deals with the future. Over time, individual preferences and technology (both affecting value weights and quantities of national income) change, and these changes are uncertain. Again, policies to be appraised affect these changes. Besides such "structural" changes there are changes connected with economic fluctuations of various amplitude and duration. These, likewise, are related to the policies to be appraised. All these problems are of interest for benefit-cost analysis, input-output studies of the Leontieff type, and other systematic attempts at a quantitative economic appraisal of policy.

Practical approximations to a solution of some of these difficulties are possible but only under restrictive assumptions with respect to institutions, preferences, technology, and time periods. Frequently, these assumptions are not made explicit when public water development projects are appraised through benefit-cost analysis and when judicial decisions and arguments before the courts involve "equity" of water alloca-

tion and the interpretation of "reasonable" and "beneficial" water use.

For policies of broader scope, the restrictive assumptions needed for benefit-cost analysis become too burdensome. In appraising such policies, it is useful to employ as criteria their effects upon significant conditions which facilitate or impede an increase of national income rather than focus on such an increase itself. This approach to policy criteria relies heavily on economic theory but less on maximization. This approach is greatly interested in economic history and in relating one time period to another but not necessarily through increasing the number of variables and equations in mathematical models. This approach is especially suited for natural resources policies and has been discussed in detail elsewhere.[33]

The emphasis of this approach is on minimum standards in resource use rather than on the optimum use; on establishing base levels rather than on locating peaks; on avoiding dead-end streets and on keeping direction rather than on computing the shortest distance; on mobility and adaptability of productive factors rather than on their optimum combination; on reducing institutional obstacles to water development rather than on maximum development; and on provisions in water law that facilitate changes over time in water allocation rather than on an optimum water allocation at particular times and places.

This approach does not pretend to establish criteria for maximizing social satisfaction. But it offers effective direction signals for pursuing the public interest turn by turn.

It becomes apparent now in what sense security and flexibility of water rights can be regarded as economic criteria and why so much attention was given to these concepts previously (Sections II, IV). Both relate to significant conditions which facilitate or impede an increase of national income in a world of persistent but uncertain change. We saw that the criterion "security" is as significant for water development as the criterion "flexibility" is for water allocation. We also saw that the logical polarity of these criteria does not necessarily imply that they are "competitive" if applied in an economically meaningful way. This is a corollary of the economic interdependence of policies concerned with water development and those concerned with water allocation (Section V).

What are the conclusions for the relations between economics and law—taken as two important social science disciplines? Economics cannot define social optima which the law—as "social engineering"—should aim to realize. What economics can do, however, is to explain why and how far certain conditions, which are decisively influenced by the law, facilitate or impede an increase of national income. Economics can point out the essential features of conflict situations and the probable consequences of changes in statutory provisions, judicial decisions, and administrative regulations. Sometimes these consequences can be shown in quantitative terms under restrictive assumptions. More often the consequences can be indicated merely in terms of direction and in terms of relative magnitudes and rates of change.

Economics need not be passive in fulfilling this function. Frequently, a conflict situation can be identified in economic terms before it has arisen in law as a controversy. After it has arisen as a controversy, the essential economic features may not be clear to the contestants themselves.

[33] S. V. Ciriacy-Wantrup, *Resource Conservation: Economics and Policies* (Berkeley: University of California Press, 1952), 395p. See especially Chapter 18.

A first, but necessary, step toward implementing such a relation between economics and law is mutual understanding with respect to the interpretation and application of key concepts used as economic criteria. In the area of water law, such concepts are security and flexibility of water rights.

To such an understanding, both normative and positive economics can make a contribution. If a value judgment is permitted, one may add that the contribution by positive economics has been far greater, and that this will probably hold also in the future. The law, on the other hand, is essentially a normative discipline. For this reason, doubt was already expressed (Section I) the "integration" of law and economics is possible or desirable. But in spite of—or possibly because of—differences in basic orientation, positive economics and law have many complementary relations. To explore and to strengthen these relations will benefit both social science disciplines.

Diminishing Returns in the Depletion of Mines

By JOHN C. MURDOCK*

A RATHER substantial amount of economic literature has been devoted to mining since Adam Smith's *Digression on Silver*. Running through this literature are several extended controversies on the relationship of various aspects of mining to some of the more general economic concepts. One such controversy concerns the significance which diminishing returns has for mining, if indeed it has any at all.[1]

With Smith and those members of the classical school who commented upon the subject there appears to have been unanimous agreement that diminishing returns is applicable to mining.[2] By this they meant that the marginal yields from mining diminished as mines of poorer quality were brought into production but, in addition, and this idea was explicitly set forth by John Stuart Mill, diminishing returns also referred to the performance of a particular mine.[3]

It is this latter application of the concept of diminishing returns to which Alfred Marshall subsequently took exception. In his *Principles* Marshall expressed the contrary view that mines do not give diminishing returns in the sense that farms do.[4] To explain his position, Marshall further refined the meaning of the performance of a mine. He distinguished between returns achieved from increasing the current rate of recovery from a mine and the yields which are achieved in successive time periods from a mine.[5] The former, he felt, cause no increase in the expenses of recovery other than those which would ultimately occur even if recovery were carried out at a slower rate by a smaller amount of capital and labor. In other words, Marshall felt that costs of production in mining are entirely neutral with respect to the rate of recovery. A greater application of capital and labor to a mine simply results in a proportionately more rapid recovery; the total returns to the factors are obtained sooner by them than they would be if a lesser quantity of capital and labor were utilized but the return per unit of factor is precisely the same regardless of the rate at which recovery takes place or the number of factors employed.

On the other hand, he admitted that as a mine is depleted it becomes increasingly difficult to recover the mineral. Under such conditions there is a tendency for the yield of the mine to diminish as it is depleted. But he observed that the yield of the mine is not a *net* yield, like the yield in the case of the law of diminishing returns. Following Ricardian rent doctrine, Marshall felt that a *net* yield, such as is received by certain agricultural land,

* Department of Economics, University of Missouri.

[1] A closely related, yet separate, subject for disagreement was the nature of the return to the landowner which resulted from the exploitation of minerals. More specifically, the question was whether royalties so obtained should be considered a rent or not. The origin of this discussion was in a Ricardian criticism of Smith. It was still the subject of disagreement as late as Marshall and Taussig. Discussions on this point are found in: D. Ricardo, *The Principles of Political Economy and Taxation*, Everyman's ed. (London: Dent, 1911), pp. 43-44, 46-47; A. Marshall, *The Principles of Economics*, 8th ed. (London: The Macmillan Company, 1920), pp. 438-439; F. W. Taussig, *Principles of Economics*, 4th ed. (New York: The Macmillan Company, 1939), Vol. II, pp. 137-140; and L. C. Gray, "Rent Under the Assumption of Exhaustibility," *Quarterly Journal of Economics*, May 1913, pp. 466-489.

[2] Cf., for examples: A. Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, Modern Library ed. (New York: The Macmillan Company, 1937), pp. 181, 213-214; Ricardo, *op. cit.*, pp. 46-47; J. S. Mill, *The Principles of Political Economy*, from 5th London ed. (New York: Appleton, 1881), Vol. I, p. 242, but cf., 583-584.

[3] Cf., Mill, *op. cit.*, p. 583.

[4] *Op. cit.*, pp. 166-167.

[5] Mill (*op. cit.*, pp. 242, 583-584) had noted both these aspects of mining but did not take pains to distinguish them explicitly.

is a payment for the "original and indestructible powers of the soil."[6] It is a constantly recurring income. The yield of a mine, however, takes the form of a royalty received by the landowner and is akin to amortization. It is a payment made for the depletion of the valuable deposit in the ground. Such a yield cannot continue indefinitely for it is composed of a part of the mine itself and once removed this part can yield no more. When this kind of return diminishes it reflects only the withdrawal of smaller quantities of the valuable mineral as the mine approaches exhaustion.[7] Thus, Marshall felt it was misleading to say that mining conforms to the law of diminishing returns.

The extremity of this position is such that, despite the great weight carried by Marshallian opinion, few students of the economics of mining accept it at the present time without qualification in some degree. Yet Marshall did contribute to the understanding of the returns from mining by distinguishing those production problems associated with the current rate of recovery from those associated with the progressive exhaustion of the mine. In so doing, however, he laid mining operations open to the claim that diminishing returns could originate from two sources instead of one.

Subsequently those writing on the topic have suggested that increases in the *rate* of recovery through the addition of capital and labor will reflect diminishing returns in the technical sense that Marshall denied.[8] From the empirical evidence cited by these writers as well as the now generally recognized nature of the law of diminishing returns it may be concluded that the law is operative for this aspect of mining.

While differing with Marshall on this point, however, there seems to be general agreement that the law of diminishing returns in the technical sense does not apply to (and hence does not explain) the tendencies toward reduced yields per unit of capital and labor employed in a mine and toward increased costs per unit of mineral recovered as a mine is depleted. No one doubts that these yields do tend to fall nor that recovery of the mineral becomes more difficult. Nor is there question as to the immediate causes of these tendencies. A variety of production problems may grow up as depletion proceeds. There is greater difficulty in reaching the deposit; its richness declines; problems relating to the pumping of water out of and fresh air into the mine are likely to increase. These and a host of other difficulties, which can be lumped together as reduced accessibility and lower quality of the deposit, explain why the yield from a mine tends to decline as depletion progresses. There is agreement on these causes and, apparently, on the belief that they are not a manifestation of the law of diminishing returns.

This unanimity of opinion, nevertheless, does not extend to the matter of explaining *why* the law of diminishing returns is not operative during the depletion process. Marshall's explanation that the return being considered is not

[6] Ricardo, *op. cit.*, p. 33.

[7] Marshall's analysis reflects his position in the rent versus royalty controversy mentioned in Footnote 1. He was of the opinion that a royalty is not a rent. It is a part of the marginal supply price of the mineral along with the expenses of operating the mine since it represents a dimunition of the mine's value (*op. cit.*, pp. 438-439).

Both Taussig and Gray challenged this position. Taussig (*op. cit.*, p. 140) pointed out that physical exhaustion of a store is not, by itself, sufficient to explain royalties. The basis for royalties is the value of the store. And Gray (*op. cit.*, pp. 481-484) observed that the value of a mineral *in situ* depends upon whether it will yield a surplus above cost of production. If it will not, he indicated, it has no value and can yield no royalty.

[8] Gray (*op. cit.*) first challenged Marshall on this point with an impressively solid argument. Later attempts have been made by others to refine Gray's criticism. Cf., particularly, W. Roberts, "Diminishing Returns in the Mining Industry," *The Journal of Land & Public Utility Economics*, February 1939, pp. 21-28.

a net return and by implication, therefore, is not subject to the same forces for its determination has not received universal support. Another explanation, suggested by Warren Roberts, appears to be based on the argument that the character of the supply of factors changes during depletion. This change, it is said, alters the point at which diminishing returns begins to influence the current rate of recovery from the mine but it does not involve diminishing returns with depletion.[9] That neither of these arguments is conclusive can be shown by a re-examination of the nature of the depletion process and of the conditions requisite to the emergence of diminishing returns.

Historically the emphasis in the analysis of mining production has been upon how the operator alters conditions through changes in the capital and labor applied to the mine. This has, of course, led to some useful results. But it is possible that this approach has led to a misunderstanding or at least to greater difficulty in analyzing some aspects of mining operations. Such is apparently the case with the depletion process.

It is true that, with variations in the current rate of recovery, the significant variable agents of production are capital and labor. At any given moment of decision on alternative rates of recovery, the mine itself is a fixed quantity. This is certainly not the case over the entire period of depletion of the mine, however. Over this longer period, the process which exhausts the mine by removing its richer deposits becomes cumulatively of more and more importance.

This exhaustion or depletion, which may take the form of various physical and engineering problems such as those mentioned earlier, represents a reduction in the quality of the land factor being employed. Alternatively, depletion can be viewed as a quantitative reduction in the natural resource being utilized in the productive operation, but the consequence is the same so far as production is concerned. To regard the quantity of land as constant under conditions of depletion simply because the surface area of the land is unchanged is superficial. From its functional standpoint, and this is the significance of any factor to production, the input of this agent has been reduced.[10] Thus, in the analysis of mine depletion, the mine itself must always be regarded as a variable factor. While capital and labor quantitities *can* be altered, the input of land is of necessity undergoing a reduction.

Such a reduction in factor input must be expected to bring a decline in total yield from the mine per unit of time unless there is a compensating increase in other factor inputs or unless a part of the previous input of the reduced factor had been redundant—a possibility for the short-term in mining. In any case, yields must eventually decline and finally cease altogether as the input of land is continually reduced.

This, of course, does not necessarily indicate the appearance of diminishing returns in any technical sense. Declines in mining yields as a result of the reduction of inputs do give a *prima facie* appearance of diminishing returns because the physical entity of the mine itself appears to be unchanged in the face

[9] Cf., Roberts, *op. cit.*, pp. 24-28. In the most recent comment which has been made on the subject, Donald Carlisle simply suggests that the law of diminishing returns does not apply to this aspect of mining. No explanation is offered. Cf., D. Carlisle, "The Economics of a Fund Resource with Particular Reference to Mining," *American Economic Review*, September 1954, pp. 595-596. This article presents the most comprehensive existing theory of mining production, however.

[10] Whether to consider the effect of depletion as a quantitative or qualitative change in factor inputs in the case of a given operation rests, I believe, with the purpose for which the concept of land is to be used. In the analysis of diminishing returns it is convenient to view these changes as quantitative since diminishing returns is customarily framed in terms of variable quantities and homogeneity of quality.

of the deline in yields. This is illusory, however.

In its technical usage diminishing returns refers to the effect on production which occurs when one factor replaces another for which it is not a perfect substitute. The law is a reflection of this relationship; it is not a matter of absolute factor quantities. In theory, at least, there is nothing to imply that the decline in inputs of land, which occurs in the course of depleting a mine, could not be accompanied by an alteration of the other factors in *due* proportions effecting a decline in output but avoiding the effect of diminishing returns to any factor input. Furthermore, the long length of the depletion period would encourage the maintenance of a roughly fixed proportion of these factors over the life of a mine since there would be sufficient time to vary more of the inputs. To this extent, therefore, the conclusion reached by Marshall is sound. There is no necessary appearance of diminishing returns simply because yields decline as depletion proceeds.

But it should be understood that circumstances can prevent the maintenance of such proportions and that, in any case, where variable proportions of the factors occur in the production process the result is a diminished return to the factor becoming relatively more abundant. If, for any reason, it is not possible to reduce the capital and labor factors proportionately to the depletion of the mine, the factor combination is altered. Capital and labor are forced to combine with land under less and less favorable circumstances. A disproportionate increase in capital and labor must be employed per unit of mineral removed because these factors cannot provide a perfect replacement for the land which has been withdrawn through production in the past.

The form this substitution (of capital and labor for land) takes may vary. It may be a straight shift to more and more capital and labor as the land factor is reduced. This occurs in mining as, for example, the face of an ore body retreats. Where the fundamental mining procedure remains the same but the problems associated with carrying it out multiply, the capital-and-labor to land ratio is increased with no substantial change in the nature of the capital and labor. The familiar experiences in coal mining as pits are deepened are illustrative of this kind of change in factor proportions.

Proportions may also be altered by way of mining technology. This is to say that in many instances exhaustion of a particular grade of ore from which the mine is equipped to produce need not mean total exhaustion of the mine. By employing different techniques of recovery it is sometimes possible to exploit a lower grade of ore successfully. The change of techniques ordinarily requires a change in the *form* of capital and labor inputs, whereby they can be more efficiently substituted for the land. In general, these techniques for mining lower grades of ore also limit flexibility in the rate of operations, depending upon large volume of processing to be economical. This in turn means they process the natural resource input more rapidly but are likely to require increased inputs of capital and labor in the bargain. Factor substitution through technological change has become more and more frequent as currently known stocks of natural resources are put under pressure of greater and greater demands. The practice currently extends, for example, from gold through crude oil production.

Regardless of the form which this substitution process takes, however, the

conditions described are clearly appropriate for most examples of mining. In most cases diminishing returns to capital and labor may be expected to develop over the course of exploitation of a mine. On the reverse side of the coin, of course, the more intense use made of the remaining natural resource input suggests that despite a downward trend in total yield the *marginal* return to the land may well rise as depletion continues.

In its usual formulation diminishing returns is stated in terms of the short-run. It is indeed a fundamental consideration in short-run production theory and, perhaps, makes its appearance most frequently in the short-run because of the limited capacity for maintaining the desired factor proportions in the shorter periods. But the short-run should not then be misinterpreted as a requisite for diminishing returns. While the period of depletion of a mine may span a longer time than is customary for short-run production considerations, the occurrence of diminishing returns is also likely to be appropriate to depletion since the ability of the producer to maintain due proportions of the factors is also apt to be limited under these circumstances. The most important difference is that in the case of short-run production the effect of diminishing returns is related only to current rates of production and current inputs. In the case of depletion the effect relates to long-run input trends and to the level of cumulative total recovery from the mine. Yet this is not a difference of substance. The crucial ingredients of the phenomenon of diminishing returns are the necessity of substituting one factor for another and the imperfection of substitutibility among those factors. In the course of depletion of a mine these conditions can be, and normally are, met.

Railroad Consolidations and the Regulation of Abandonments

By MICHAEL CONANT*

THE overinvestment in American railroads during the period of local monopolies has been bearing bitter fruit. Although earnings on total railroad investment are estimated to be only slightly below other sectors of domestic industry, the outlook is not bright.[1] Increasing competition of rival forms of transport and constantly rising costs have created a growing incentive for the consolidation or joint operation of parallel main lines and roads.[2] The full benefit of such consolidations or agreements can be had only by disinvestment in many present main lines through abandonment. The purpose of this paper is to review current abandonment practices and procedures and to suggest supplementary statutes to facilitate main line abandonments.

Competition of other transportation facilities has been found to be the largest single cause of operating losses leading to applications to abandon railroad lines.[3] Highway transport has made the largest inroads into freight revenues. Highway competition and airline competition have made for an even greater decline in passenger revenues.[4] General rate increases do not offer an easy solution to the problem in this era of increased rivalries.[5] Constantly rising railroad maintenance costs accentuate the problem of light traffic lines and segments.[6] The result is operation with much greater excess capacity that is necessary for efficiency.[7] It is the pressure on railroad managements to reduce this excess capacity by more intensive use of fixed plant[8] that has revitalized the consolidation movement. The ultimate cost savings will depend on the railroads' ability to abandon longer rail segments than in most previous abandonment cases. It is for this reason that supplementary abandonment procedures are needed.

* University of California, Berkeley.

[1] See comment in *Railway Age*, December 26, 1955, p. 8, where the net return on investment in American railroads is estimated at 4.15 percent in 1955 and 3.28 percent in 1954.

[2] Two major mergers of parallel roads now under consideration are the Chicago and North Western with the Chicago, Milwaukee, St. Paul and Pacific and the Northern Pacific with the Great Northern and two connecting roads. *Railway Age.*, November 21, 1955, p. 13, and January 30, 1956, p. 13.

[3] U. S., Interstate Commerce Commission, Bureau of Transport Economics and Statistics, *Railroad Abandonments 1920-1943* (Washington, D. C.: 1945), pp. 19-26.

[4] U. S., Congress, Senate, *Railroad Passenger Deficit Problem*, (84th Congress, Senate Document 24, 1954).

[5] Although the railroads have received a number of general rate increases, they are currently requesting relief from part of the regulatory controls so they can engage in price-cutting rate competition. See Jervis Langdon, Jr., "The Regulation of Competitive Business Forces: the Obstacle Race in Transportation," *Cornell Law Quarterly*, Fall 1955, p. 57; Earnest W. Williams, "Railroad Rate Levels and Earning Power in an Era of Competitive Transport," *Land Economics*, November 1949, p. 405. Cf., Joseph B. Eastman, "The Adjustment of Rates between Competing Forms of Transportation," *American Economic Review*, Supplement, March 1940, p. 124.

[6] One of the presently-contemplated main line railroad mergers will illustrate the type of expenses to be saved by abandoning one of two parallel lines. The average cost of maintenance of way and structures per equated track mile on the Chicago, Milwaukee, St. Paul and Pacific in 1954 was $3128, and that on the Chicago and North Western in the same year was $2997. *Moody's Transportation Manual*, 1955, pp. 43 and 131.

[7] The interrelations of overinvestment and excess capacity are explained in J. M. Cassels, "Excess Capacity and Monopolistic Competition," *Quarterly Journal of Economics*, May 1937, p. 426 at 440-443. It must be noted that the Weeks Report demands substantial excess railroad capacity for defense purposes. A question arises as to whether private business should bear the cost of public defense. Presidential Advisory Committee on Transport Policy and Organization, *Revision of Federal Transportation Policy* (Washington, D. C.: 1955), p. 4.

[8] See Herbert Ashton, "Railroad Costs in Relation to the Volume of Traffic," *American Economic Review*, June 1940, p. 324; Harold Hotelling, "The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates," *Econometrica*, June 1938, p. 242 at 264.

TABLE I—ABANDONMENT DECISIONS, NUMBER PROTESTED, NUMBER DENIED, AND CASES DISMISSED, 1946-1955

Year	Abandonment Decisions	Number Protested	Percent Protested	Number Denied	Percent Denied	Cases Dismissed Without Decision
1946.....	43	17	39.5	6	14.0	4
1947.....	64	14	21.9	1	1.6	1
1948.....	63	22	34.9	6	9.5	9
1949.....	59	26	44.1	3	5.1	3
1950.....	88	34	38.6	8	9.1	7
1951.....	63	32	50.8	5	7.9	5
1952.....	88	34	38.6	4	4.5	5
1953.....	78	20	25.6	1	1.3	5
1954.....	69	20	29.0	3	4.3	6
1955.....	62	11	17.7	0	0	5
Total.....	677	230	34.0	37	5.5	50

Source: U. S. Interstate Commerce Commission, *Annual Reports*, 1946 to 1955.

Number and Size of Abandonments

The volume and character of abandonments before the Commission in the postwar period generally follows the pattern of the previous 25 years.[9] During the ten years, 1946-1955, 734 applications for abandonment were filed. This was an average of 73.4 per year. During the 25 year period, 1921-1945, 2363 applications for abandonment were filed, an average of 94.5 per year.[10] The smaller postwar number is probably due to the high demand for rail services and increased earnings of this period.

As shown in Table I, the Commission rendered 677 decisions in abandonment proceedings from 1946 to 1955. Of these decisions, 230 or 34 percent were protested. Only 37 of the decisions resulted in denials, being 5.5 percent of the total decisions.[11] In addition, there were 50 cases in the period which were dismissed without decision. Professor Cherington's comments on the similar low ratio of denials to decisions in the earlier period is applicable to the postwar period. He wrote.:

"Superficially these statistics would seem to indicate that the Commission has acceded to the wishes of the railroads and permitted abandonment in the great majority of cases. When an application is filed, the chances could appear to be nearly sixteen to one that a certificate would be granted. From such a record one might conclude that the Commission has been liberal indeed in meeting the wishes of the applicants, and that relatively little controversy exists in the ascertainment of what the public convenience and necessity permits. Such a conclusion, however, would be entirely misleading. The formal denials of certificates of public convenience and necessity have been relatively few in number but they have been of great importance in curbing the natural trend toward railroad abandonments. Particularly in the early cases where certificates were denied, the Commission indicated in plain and unmistakable language the comparatively narrow limits and conditions within which it would permit abandonment to take place. The criteria laid down in these early cases and resorted to again in the later denials have

[9] Interstate Commerce Commission, Bureau of Transport Economics and Statistics, *Railroad Abandonments 1920-1943* (Washington, D. C.: 1945), p. 7; Charles R. Cherington, *Regulation of Railroad Abandonments* (Cambridge: Harvard University Press, 1948), p. 100.

[10] *Ibid.*

[11] From 1920 to 1943 the rate of denials was 6.0 percent. Allowance for denials which were followed by authorization in a subsequent year lowers the rate to 3.1 percent. Interstate Commerce Commission, *op. cit.*, p. 7.

served as warnings to prospective applicants. In consequence, many segments of railroad, particularly unprofitable branch lines of otherwise profitable systems, have been continued in operation . . . ."[12]

Many of the cases in which no protests were filed involved short segments of road for which there were adequate substitute lines or for which the need had disappeared. The important policy questions arose in the protested cases. Although the denials were only 5.5 percent of the total decisions, they constituted 16.9 percent of the total protested cases.

The majority of cases in which protests led to denials involved longer segments than those in which abandonment was granted. As shown in Table II, abandonments granted from 1946 to 1955 covered 9316 miles or 14.6 miles per abandonment. The 36 denials (the one ferry operation excluded) of 1650 miles averaged 45.8 miles per denial. The longer the segment, the more likelihood that entire villages or towns will be left without alternative rail service. The local coal dealer, building supplies dealer, and elevator operator, who may be the only users of the line for carload freight, are almost sure to file protests. To the carrier, these longer segments are the ones causing the greatest losses. Consequently, it is these segments that bring the hardest fought controversies before the Commission.

TABLE II—MILES OF ROAD AND MILES OF OPERATIONS UNDER TRACKAGE RIGHTS ABANDONED, 1946-1955

Year	Abandonments Granted			Abandonments Denied		
	Total Miles	Miles of Road	Miles of Operations Under Trackage Rights	Total Miles	Miles of Road	Miles of Operations Under Trackage Rights
1946.....	670	640	30	273	273	—
1947.....	1241	1175	66	2	2	—
1948.....	907	660	247	48	48	—
1949.....	1185	873	312	29	29	—
1950.....	955	867	88	110	110	—
1951.....	564	544	20	181	181	—
1952.....	1306	1074	232	99	99	—
1953.....	1102	795	303	6	6	—
1954.....	873	848	25	14	14	—
1955.....	513	429	84	0	0	—
Total.....	9316	7909	1407	1650	1650	

Source: U. S. Interstate Commerce Commission, *Annual Reports*, 1946 to 1955.

Public Convenience and Necessity

The Transportation Act of 1920 gave the Interstate Commerce Commission regulatory power over railroad abandonments. It made the sole test in judging abandonments that they be consistent with public convenience and necessity.[13] This highly amorphous standard vested great discretion in the commission. Only the decisions of the federal appellate courts set the statutory and constitutional

[12] Cherington, *op. cit.*, p. 102.

[13] 49 U.S.C. 1(18); *Colorado* v. *United States*, 271 U.S. 153, 168 (1926). This presumes that regulatory commissions will allow rate increases to protect the carrier's total earning levels when one segment shows losses.

limits to this discretion. The result has been a long series of cases, impossible of classification, developing no systematic detailed criteria on which later applicants can base their petitions.

The Interstate Commerce Commission abandonment decisions are said to involve a balancing of interests.[14] Justice Brandeis explained the burden on interstate commerce of the losing branch, that significant losses in one division of a railroad must be paid for by the shippers in other divisions.[15] On the other side are the shippers of bulk products, such as grain, coal, sand and stone, who have invested in fixed plant in reliance on continued rail service. They stand to lose the greater part of their investment if the line is abandoned because truck shipment of their products is too costly for them to compete. The commission decision involves the weighing of these incomparables, the burden on interstate commerce against the estimated losses to those shippers who cannot operate profitably without rail transport.

The planning of long-run investment and of consolidations are seriously impeded by present abandonment procedures. The only small group of abandonment applications before the Interstate Commerce Commission which have estimable results are those to abandon the entire line of a railroad. If the firm has present operating losses and the prospect is that deficits will continue, the commission must permit abandonment. The constitution has been interpreted to forbid forcing an entire business to continue operating at a loss.[16]

The more usual and more complicated cases are those proposing abandonment of one segment by an applicant whose overall system shows net profits. Here the applicant may be ordered to continue operating the segment at a deficit if the commission finds that applicant's loss is more than offset by the public need for that rail service. The courts will not reverse such a finding unless it is shown to be unsupported by evidence.[17] The Interstate Commerce Commission is thus given great discretion in the relative weight given the factors bearing on the application.

The first burden of proof is on the carrier to establish a recent history of operating losses on the segment concerned. Failure to prove these losses is ground to deny the application.[18] The commission will also give weight to the projected out-of-pocket cost of rehabilitating the line due to deferred maintenance or flood damage.[19] From these data the savings from the proposed abandonment can be estimated. The public need for continued service is conjectured from testimony of the protestants as to the inconvenience and losses they maintain the abandonment will cause them. Each case is determined on its own facts, and no minimum volume of traffic per mile has been established to create a presumption in favor of the need for continued service. The nature of the shipments and the alternative costs of transport are considered. The total investment loss to shippers and the probable economic growth of the region are also taken into account. Upon reaching a decision, the commission issues an opinion setting forth the relevant facts which support its

[14] See discussion in Cherington, *op. cit.*, 125-136.

[15] *Colorado* v. *United States*, 271 U.S. 153, 160 (1926). For this reason, the Interstate Commerce Commission has jurisdiction over abandonments of intrastate branches of interstate carriers. *Ibid.*, 167-168.

[16] *Brooks-Scanlon Co.* v. *Railroad Commission of Louisiana*, 251 U.S. 396 (1920). Recent examples are the 364-mile line of the *Missouri and Arkansas Ry. Co. Abandonment*, 271 I.C.C. 171 (1948), and the 173-mile line of the *Texas Electric Ry. Co. Abandonment*, 271 I.C.C. 391 (1948).

[17] *Virginian Ry. Co.* v. *United States*, 272 U.S. 658, 665 (1926).

[18] *Great Northern Ry. Co. Abandonment*, 271 I.C.C. 207, 215 (1948).

[19] *Chicago, B. & Q. R. Co. Abandonment*, 271 I.C.C. 261, 268 (1948).

findings. However, it generally does not indicate the relative importance of the different factors or which of them is controlling.[20] Hence, each applicant faces the commission with the same great uncertainties as did previous applicants. No detailed body of rules is built up on which carriers can base their planning of abandonments. As a result, the economic basis of consolidation and coordination—the rational allocation of investment—is undermined with this uncertainty that unprofitable plant may have to be kept in use.

Conclusions and Proposals

The uncertainties of the present railroad abandonment statute and procedures make them inadequate as a necessary part of planning a rational railroad system. Under present procedures, shippers of bulk products have valid objections that they have made fixed investment, perhaps recently, in reasonable reliance on the continued existence of railroad service. Carriers validly object that they cannot plan abandonments under present procedures because Interstate Commerce Commission decisions are in most cases completely unpredictable. For example, carriers' accounting methods showing losses may be rejected in favor of accounting methods which will show that the branch or segment was profitable. This usually leads to a denial of an abandonment application. The present Interstate Commerce Commission rule as to depreciation accounting is illustrative. In an abandonment application, a carrier may not write on as an expense the estimated wearing out of roadway structures.[21] Instead the expense accounts may show only out-of-pocket expenditures on replacement. Thus, if in contemplation of abandonment, a carrier makes no replacement expenditures on a line in a year, it may cause its accounts to show a profit although real costs of operation would show losses.

Current procedures will actually impede any movement to large-scale consolidations and abandonments designed to decrease overinvestments in railroad lines. In the consolidation of parallel main lines, carriers will usually be unable to show a history of losses on the line to be abandoned. Because each main line will have had a large amount of bridge or through traffic, each will have a record of profits.[22] In order to abandon one of the lines under present procedures, consolidating lines would first have to shift all bridge traffic to one line. Then the consolidated carrier would have to operate the line it wished to abandon for local traffic only for a period of years to establish a history of losses. Even then, the Interstate Commerce Commission might deny the abandonment on the ground that these losses are overbalanced by the public need for the railroad service.

The benefits of main line consolidations can be had only by a new, supplementary law of abandonments. The first requisite is that there be prior scheduling of abandonments. A period of notice fixed in each case by the Interstate Commerce Commission should precede a major abandonment. Only with adequate notice can shippers and carriers plan alternative routes and conserve and redirect investment accordingly. The second requisite is that carriers be given an absolute right of

[20] Isaiah L. Sharfman, *The Interstate Commerce Commission* (New York: The Commonwealth Fund, 1935), Vol. III A, p. 338.

[21] *Chicago & N. W. Ry. Co. Abandonment*, 275 I.C.C. 759, 772 (1951); *Chicago & N. W. Ry. Co. Abandonment*, 282 I.C.C. 525, 530 (1952).

[22] Current abandonment applications must compute past revenues by allocating to the segment to be abandoned all the bridge traffic which has moved over it, even though the bridge traffic is about to be shifted to a more efficient parallel line. *Chicago & N. W. Ry. Co. Abandonment*, Finance Docket 16868 (1952); *Chicago, Burlington and Quincy R. Co. Abandonment*, 271 I.C.C. 261, 274 (1948).

abandonment, subject to the notice requirement and certain other possible conditions. Long-run investment planning requires a degree of certainty that commission veto precludes the present procedures from having. However, the right of abandonment in the new statute could be circumscribed by conditions for adequate protection of shippers without putting a power in the commission to prevent the abandonment entirely. Types of conditions in abandonments are discussed below.

Under the new scheduled abandonments statute, the carrier would begin by filing an application for hearing with the Interstate Commerce Commission. The application would indicate the line or segment to be abandoned and suggest a proposed period of public notice before the abandonment would be effective. If no protest were filed within the protest period, the application would be automatically approved. Then the carrier would be obliged to post public notice in all terminals for the entire notice period. At the end of the notice period the line would be abandoned.

If protests were filed, the Interstate Commerce Commission would set the application down for hearing. Since the right to abandon would be absolute, the only question in contest would be the length of the appropriate notice period. At the hearing, shippers could present evidence as to the nature of their products which must be shipped by rail, the amount and depreciation rates of their fixed plant, the sale value of their plant to other firms not requiring rail transport, and the cost of dismantling and moving fixed plant. The carrier could present its estimates as to these factors and offer plans to assist in relocating plants of protestants. After considering all these factors, the Interstate Commerce Commission would set a period of public notice to elapse before abandonment. The statute might put an upper limit on the notice period of, say, seven years. In considering the new type of abandonment application, no questions of costs or revenues of the carrier would be material. The right to abandon, being absolute, would in no way depend upon past losses or profits.

The Interstate Commerce Commission should be given power to raise rates on any segment or branch during the period of public notice pending abandonment. It would probably use such power quite sparingly where one of two parallel main lines was to be abandoned, as it would create inequities between competing shippers on the two parallel lines. However, on branch lines higher rates are in many cases a workable alternative to abandonment. On many branch lines a few shippers have a vested interest in rail service only because of the unrealistically low rate structure imposed by government in the attempt to equalize geographical advantage. On such lines rate increases would correct inequities by making rates more consistent with costs. Such increases allowed during the period of public notice pending abandonment might cause the carrier in some cases to withdraw its abandonment application.

The new procedures, creating an abandonment of right, would not replace present permissive abandonments. The present procedures, while inadequate for larger scale abandonments, are nevertheless satisfactory for most cases of branch lines with an uncontestable history of losses. Hence, the new statute should not replace the present one, but merely supplement it. No conflict would arise between the two distinct procedures. A carrier would have a choice of filing initially under the old procedures or the new. If it filed first under the old procedures and immediate abandonment

was denied, it could then file under the new scheduled abandonments statute. If it filed first under the new statute and a long notice period was set, it could file an interim petition under the old statute before the notice ran out, requesting permission to abandon earlier because of sharp traffic declines causing losses.

Conditions precedent to abandonment are perhaps even more appropriate under the proposed new abandonment of right than under the old system. Under the new procedures, shippers and other protestants would be precluded from arguing in the commission against the abandonment itself. They should, however, be given the possibility of easing their adjustment to this technological change by arguing for conditions precedent to the abandonment. Many types of conditions applied in past cases under the old statute might be used. The most frequently used one under the old statute was that the carrier offer to sell the segment or line, usually at scrap value, to anyone who would purchase it for continued operation.[23] Conditions protecting the seniority of employees and ordering compensation at previous wage rates though workers must take less skilled jobs are common.[24] Less frequently, a carrier may be required to build new connecting lines to a shipper's place of business.[25]

More difficult problems arise as to conditions precedent requiring the railroad to pay compensation to private individuals. It is debatable whether this involves a constitutional problem of taking property without just compensation. The carrier does have the alternative of continuing to operate the segment of line and foregoing abandonment. Under the present permissive abandonment statute, Congress has not given the commission power to attach such conditions.[26]

A strong argument can be made that firms which have invested in reasonable reliance on the continued existence of a rail line should not be forced to lose the greater part of their investment. The remedy would be to condition the abandonment of such line on the carrier's compensating bulk shippers for their investment loss. Such conditions would be available only to the types of businesses which cannot compete in their markets by shifting to truck transport. Grain elevators and shippers of coal, sand, stone and clay products are in this class. The carrier could be required as a condition precedent to abandonment to pay these firms the difference between the net depreciated value of their plants and the amount realized by sale when the rail line is abandoned. In other cases, the carrier might be required to move or pay the cost of moving dissembled plant and equipment from the abandoned line to another nearby line.

The final suggested statutory change is an alternative protection for shippers who would be unable to carry on without rail service. This is special tax relief. Accelerated depreciation allowances could be given on plant not convertible to other uses. The Internal Revenue Code would have to be amended to allow a shortened depreciation period. In each case the shipper would be allowed to adopt a depreciation period on plant equal to the period of notice required for abandonment by the railroad. These tax revisions would greatly facilitate the operation of the proposed abandonment procedures

[23] Cherington, *op. cit.*, 172-174.

[24] *Chicago, B & Q. R. Co. Abandonment*, 257 I.C.C. 700 (1944); *Interstate Commerce Commission* v. *Railway Labor Executives Association*, 315 U.S. 373 (1942).

[25] *Lehigh Valley R. Co. Abandonment*, 290 I.C.C. 427, 432 (1954).

[26] Cherington, *op. cit.*, p. 178.

by allowing the disrupted shippers to spread the cost of readjustment over a number of income periods.

In considering the procedures suggested herein, one must keep in mind that the regulation of railroad abandonments is interdependent with the regulation of railroad rates. The more flexible rate structures recommended by the Weeks Report,[17] if enacted, might raise railroad revenue substantially and thus lessen the carriers' pressures for consolidations and abandonments. However, consolidations and abandonments are problems of long-run resource allocation. The sooner the carriers correct the misallocations of resources, the sooner all shippers can gain by lower rates.

[17] Presidential Advisiory Committee on Transport Policy and Organization, *op. cit.*, pp. 8-11.

Credit in an Expanding Economy: Should We Have Hundred Percent Agricultural Credit?

By CLYDE MITCHELL*

Summary of the Argument

ECONOMIC progress in welfare terms (goods and services) is assumed to be the goal toward which social planning is directed. The United States has expressed, in the Employment Act of 1946, the intention to pursue a course of economic progress in an expanding economy. Such progress will continue to result, as it has in the past, mainly from the association of more (and more efficient) capital equipment with the factors of labor and management. Productive credit assists in bringing about that association (of more capital equipment with the labor and management factors). Availability and use of credit which facilitates the creation of more capital equipment is therefore *a condition of progress.* Serious deficiency in credit-availability to various people engaged in agriculture is one important factor standing in the way of efficient production. If the United States is to maintain an expanding economy successfully, these deficiencies must be made up rather rapidly. Tentative suggestions are made in this article that new and different methods of supplying credit to agricultural producers will be needed in the next few years. These methods at first glance appear to be "radically" different from those employed by agricultural credit institutions, particularly before 1933. They are different from those envisioned in traditional economic theory which frowns on capital-goods accretion in the absence of prior money-saving. However, a closer examination indicates that with regard to "capital-goods-formation": (a) the areas of the American economy which have made the most progress have benefited from considerable "cultural fiat and social action" with regard to production credit, and (b) the traditional theory of capital-goods formation contains basic logical faults and probably never deserved the adherence of economists in the first place. In short, it is possible that these suggestions are realistic rather than radical and involve only the extension to agriculture of ideas long accepted in industrial production. More rapid progress in the field of agricultural capital formation will probably result from *social action programs* additional to and of a more comprehensive nature than have been tried in the past twenty years. Methods should be found to establish competent farm producers in a well-equipped productive operation *at the time in their lives at which it is most likely that they will be able to produce effciently.* If plans along the lines of these suggestions are put into effect, they will change the nature of the obligations which the farm producer owes to the rest of the community. A tentative exploration is made in this article into the nature of these changes.

A Short Excursion Into Traditional Ideas of Capital-Goods Formation

A. *Robinson Crusoe and His Fish Net.* The earliest economic thinkers were impressed with the dramatic way in which division of labor and specialization could increase the production of any group of workers, beyond the amount the workers might contrive without specialization. These theorists recog-

* Food and Agriculture Organization, United Nations Regional Office, Mexico City.

nized the influence of capital goods upon increased productive efficiency and correctly reasoned that an increase in the production of capital goods was a necessary condition of economic progress. For various reasons the fathers of economic thought devoted far less attention to the technological conditions of capital-goods creation than they did to "economic" conditions, rather narrowly defined. In the famous story of Robinson Crusoe, who built a fish net to increase his haul of fish beyond the amounts he could catch with his bare hands, theory took what is perhaps a wrong turn. In order to feed himself while he spent two or three weeks weaving the net, Crusoe first needed a supply of food. He saved berries. *Saving* thus appeared to the theorists to be necessary *prior* to the construction of capital goods.

B. *Capital-Goods Formation Limited By Savings.* From this interpretation of fishing technology grew the idea that capital-goods formation is limited by money savings. Basic to the theory of capital-goods formation are the assumptions of the logical system in which capital-goods formation is only one part—the "laissez-faire system," in which prices serve as the directing force for economic decisions and bring about both efficiency in production and equity in distribution of the products of man's work. These assumptions can be summarized in the phrase, "perfect competition in a perfect market," and include, subsidiarily, mobility of factors of production and the "economic man." Given these assumptions, full and efficient employment of all factors of production would be assured "as if by an unseen hand." For society to make progress, new capital goods needed to be introduced into the system. Such introductions could be made only by those who could save money. Capital-goods formation was therefore conceived to be limited by money-saving. Money-savers were changed from the usurious devils of a slightly earlier age into benefactors of society, by the writings of Adam Smith and his followers.

C. *Forced Savings.* If money means benefaction, then could a ruler, by printing a great deal of money, become a great benefactor? For a long time the people in charge of printing paper money have been intrigued by the tremendous power in the finger with which they push the starting button of the printing press. It appeared that at a motion of this finger they could bring into being great warships, buildings, dams, highways, and national monuments. But simple intelligence convinced almost everyone that such magic could not possibly be true—that these impressive accomplishments were the product of artisans and laborers and engineers rather than the button-pushers in the print shop. In fact, the button-pushers, toiling not and sweating not, were deemed to be a rather irresponsible crew in aspiring to perform magic feats. Economic logicians took pains to point out the danger of letting the printing press operators direct such important human activities as calling forth warships and buildings. Given the assumptions of the economic system which the theoreticians believed described our world, of course the printing press operators were positively dangerous. Although they might print money which called forth the construction of capital goods, their action took the entire matter of saving out of the hands of those highly-respected members of society who could save, and forced everyone, particularly the poor people, to *save* whether they wanted to or not, or whether they could spare anything from their meager existence or not. The printing-press money forced "savings" by pushing prices up, particularly of the things that

the poor people have to buy. This early discovery that money-printing might get new industries built was therefore never given adequate study because it was almost from the start believed to be irresponsible and sinful.

D. *Capital-Goods Formation in an Underemployed Economy.* However, the theoreticians discovered that in the real world, money can sometimes be printed and put into circulation to build new capital equipment without raising prices or forcing anyone to save. This can happen whenever there are resources which are not being fully utilized in the economy. If the amount of under-utilized resources is large, governments can print large amounts of money, or credit-creating institutions can create large amounts of credit, and large amounts of new capital goods can be built using the slack resources.

The admission by present-day economists educated in the classical tradition that it is *possible* to bring about the creation of new capital goods by social action (printing money or expanding credit) without prior money-saving by capitalists and without forced-saving by consumers generally, points out a serious limitation in the usefulness of traditional theory. It constitutes an admission that society, acting through laws and other institutional factors, *can* direct our economy and do it well. Society can do it better in the real world, from a goods and services standpoint, than the automatic and impersonal forces of price and competition which (by the theoreticians at least) have been depended upon for two hundred years. The theoreticians excuse themselves by admitting that the real world exhibits underemployment of resources, which the theoretical world ruled out. However, a few modern economists are re-examining the original idea and ask if the building of the first fishnet did not itself require underemployed resources. How did Robinson Crusoe manage to store up enough berries for his three weeks of net-making? He must have lived in a surplus-producing area—a partially underemployed economy. *In a society fulfilling rigorously the assumptions of the classical theory, it is entirely possible that there never could have been any capital-goods creation.* It is probable that in every society, everywhere, enough underemployed resources exist (or can be freed by adoption of new techniques) to allow great amounts of progress through social re-direction of resources.[1]

E. *Capital Formation by Social Dictate.* Whether or not we believe that the theory of capital-goods formation through prior-savings was faulty from its beginning, most economists today acknowledge that *society* quite properly engages in the process of *dictating* a great part of the capital-goods formation that now takes place. Whenever our society makes the decision that certain things must be built imperatively, *those things are built,* whether or not anyone had previously saved enough money to build them. The wartime expansion of our nation's capital equipment is an excellent example, of course. The doubling of capital equipment in the period 1939-1953 has occurred mainly because of *social direction.* That social direction included (in addition to the creations of funds) the allocation of scarce materials, government construction of plants, guaranteed or supported prices, preferential tax treat-

[1] The theoretical dilemma mentioned in this paragraph is only one of many met in attempting to move from the static "classical" theory of value to a theory capable of describing an actual economic system existing in time. The author believes that a realistic theory of economic life is possible only when theorists recognize that considerable underemployment of resources is an unavoidable matter in any machine society, and that society therefore always has the choice of making decisions which will decrease (or increase, if it wishes) the extent of under-utilization in any particular field.

ment, and many other similar measures. An exact measurement is impossible of the extent to which America's capital goods have been increased because of direct and indirect social action. The chief economist with a large American corporation argued that I was wrong—in an article I wrote in 1953 in which I said that "more than half, and perhaps almost all" of America's doubling of capital goods has occurred because of social action. He conceded that the war plants had been built with Reconstruction Finance Corporation and other direct federal money but concluded that the balance, much more than half, was expansion from *private funds*. But that misses the point! *Those private funds, profits of American business, were as large as they were because of definite decisions made by the American people, particularly the decision we made long ago not to force the American economy to conform to the requirements of traditional economic theory.*

The decision to fight the war and to build war-related industry was a social decision. Once that decision had been made, most profits became automatically guaranteed for some years to come, not only in the war industries but also in all the less- and non-essential industries. Practically all of these industries enjoyed the most tremendous prosperity they had ever known. Savings from the net profits of private corporations *did* of course finance a great deal of the growth, but most of these net profits resulted directly from social decisions completely outside of the realm of a society governed by the *laissez-faire* doctrine. There can be no doubt that a great deal of such net profits resulted from the existence of patents, trade-marks, price-fixing, and other modifications of pure competition which society has decreed or in which it acquiesced. There can be no doubt that rapid tax-write-offs, coupled with the fact that the government directly influences about one-fourth of the total income flow in the nation, now guarantee business stability at a high-profit level throughout much of the so-called "private" enterprise sector.

I should like to repeat my 1953 statement to which the afore-mentioned business economist objected: "The doubling of capital equipment that has taken place in the past fourteen years has occurred very greatly (more than half, and perhaps almost all) because of the creation of funds beyond the amounts saved by capitalists, and certainly beyond the amounts capitalists could have saved had our economy been competitive in the classical sense." In other words, the funds for capital-goods creation were funds created by society, or allowed to be created because society has not thought it wise to force business to be classically competitive. The effect of this great increase in capital equipment has been a tremendous increase in physical productivity, in goods and services, of the American economy. The results are undoubtedly better, *in the physical-productivity sense*, than a purely competitive society could have achieved.

The Changing Ratio of Capital Value to Labor in the Twentieth Century

During the course of the industrial development of modern society, the money investment in capital equipment per worker has of course increased greatly. The average cost of capital equipment associated with each worker in American industry is more than $10,000; it is almost twice this amount in railroads and utilities. Investment in capital goods per worker has increased greatly in the past half-century as the size and complexity of industrial operations have increased. The worlds of finance and industry have long been organized in such a way as to

provide these large amounts of capital equipment *without* requiring either the industrial laborer or the entrepreneur to make prior savings of large amounts of money funds. Indeed it is almost an axiom of business that new industrial enterprises be started with little or no money. (The entrepreneurs are expected to have production ability, but even that is not necessary—engineers can be hired.) *Promotional ability* is perhaps the main requisite to starting industrial enterprises, and on so precarious a basis (in the technological sense) funds are raised from investors. The promotors usually receive no-value common stock for their promotional efforts; the cost of physical plant and working funds are supplied by investors in preferred stocks and bonds. Competent studies, such as were made by Berle and Means and others, have shown that complete control, i.e. ownership of all common stocks, of America's largest industries was initially achieved with an investment of only about 7% of the real construction cost of the industries—the other 93% was furnished by investors who received securities bearing little or no right of control over the industries.

Great physical performance of the American industrial system has characterized past years and profitable financial performance has characterized most of them. These two factors have adequately justified the optimistic hopes of an institutional system that permits entrepreneurs and laborers with ideas and abilities, but without money, to associate themselves with thousands of dollars worth of capital equipment during the best and most productive periods of their lives. To take the different course suggested by Robinson Crusoe economics would be unthinkable—to require an industrial entrepreneur to work up the ladder from a common laborer to a skilled laborer to a small backyard shop to a larger shop to a small factory to a larger factory—buying the more expensive equipment in each case from the money savings he had made by abstaining from spending part of his income in the prior stage. He would be senile before he had saved the price of one forging hammer. America's Horatio Alger folklore to the contrary notwithstanding, that is not the way an industrial economy makes progress.

Progress is made because society has made a complex chain of decisions, some legal, some institutional, which bring entrepreneurs and workmen together to work, during the most productive period of their lives, with capital goods which society has decreed shall be created.

Modern societies have learned, though most elementary economics texts avoid this fact, that economic growth is self-financing and must be. In modern societies there are always highly flexible elements of under-employment of resources, even in time of greatest emergency pressures toward so-called full employment. It is the more effective employment of these factors that produces the economic growth. It is likewise true that an optimistic direction of the economy increases prices of existing property, particularly real estate. However, the creation of new instruments of money in a technologically-able economy ordinarily is matched by an influx into the economy of more goods, about as rapidly as the new money can be spent. In such countries the amount of inflation, as customarily measured, is far less than economic theorists warn against.

Of course a recognition of this does not mean that governments or banks should create money carelessly. Countries less technologically-able than the United States must be particularly care-

ful. The process must be well-planned, and the money must be funnelled into the "right" places. If the new money is not matched by real physical growth and productivity-increase wherever its influence is felt, some price inflation may need to be dealt with by direct controls. However, the generalization is a safe one that *our technological ability to increase our capital equipment (and therefore our productivity) makes it possible for any nation to afford the increases.* This is the exact reverse of the teaching of traditional economics.[2]

Acceptance of this more modern way of looking at the problem of progress and growth underlies the Employment Act of 1946. That Act expresses with the highest ceremony possible in our society, formal act of Congress signed by the President, and implemented by a top-rank professional staff, that we have to a large extent adopted a new theory of economic development, that we as a sovereign nation will do whatever is necessary to maintain an expanding, growing economy.

Capital Goods for American Agriculture in the Future

If this is a realistic picture of the changed and more accurate explanation of the tremendous technological progress made in the western world, and in the United States in particular, to what extent has agriculture shared in these changed concepts? As compared with industry, agriculture has shared very little. This generalization is not intended to belittle the great significance of the social action programs of the last twenty years. Price-supports, incentive payments, farm credit at reasonable rates, supervised credit of FSA-FHA, road-building and the REA's—all of these have helped many farm producers to adopt new methods, invest in new capital equipment, and greatly increase their productivity. However, as between in-industry and agriculture, with regard to their ability to do an adequate job of meeting the challenge of an expanding economy for the next two or three decades, the gap is still great. In considering how to improve the physical efficiency and productivity of America's farms in the future, I should like to direct attention to some new methods of expanding the amount of capital equipment available to farmers.[3]

Agriculture is the main residual area in which the "save-first-before-building" idea prevails. Twenty years of social and political concern with "the farm problem" has partially changed this idea but, in the main, too many people still think that there is an agricultural ladder and that the task of climbing it need not be materially eased by government.

I submit that the same general type of technological and financial revolution that has brought our industrial society to the place it is today needs to take place in agriculture. On the most successful farms the physical revolution has already taken place. The farm that can support a family in decent middle-class living in the midwest now requires an investment of more than $40,000. About the maximum that a farmer can borrow for such an enterprise, without parental or other family assistance, is fifty percent. Raising $20,000, or even $10,000, is a difficult matter for most American farmers. For the young farmer it is usually impossible.

[2] These ideas are a brief summary of the thesis expressed by my friend and teacher, Professor C. E. Ayres, University of Texas. They are outlined in greater detail in his book, *The Industrial Economy* (New York: Houghton-Mifflin, 1952).

[3] It is obvious in this connection that the author accepts the viewpoint that agricultural production should be increased rather than decreased—that there are no surpluses in the human sense, but only in the economic sense; and that we must learn to utilize America's abundance for humanity's still unfilled needs rather than attempt (so far unsuccessfully) to destroy it.

To improve agricultural productive efficiency for the America of the future will require that methods be found to enable competent farm producers to associate themselves with adequate capital equipment early in life, when their vigor and ambition are highest.

As in the case of industrial America, the association should prove successful. A farmer who can thus associate himself should ordinarily achieve physical productivity high enough to pay his initial loan off within about twenty years. In such cases the increased productivity has amply justified the loan. If, however, the prevailing farm-finance pattern of today is continued, those twenty years of highest physical ability must often be partially wasted on inadequate and ill-equipped farm enterprises.

If Society Takes a Hand

Society has taken a direct hand in agricultural capital formation only in a timid manner, as in the Farmers Home Administration program, and to the extent that price and income supports have had an indirect effect. American planners should consider whether or not in the process of making capital funds available to increase farming efficiency in an expanding economy, society should participate to an extent beyond anything contemplated in present laws and institutions. To this end capital funds up to ninety or one hundred precent may need to be supplied to farmers who give evidence of being good entrepreneurs. Implicit in this proposal is, of course, the proposal that some measure of competence be devised and applied to applicants for funds.[4]

Much of the new investment funds thus made available would be from sources "outside" of agriculture. Whether the loans come from private financial institutions, private institutions with government guarantees, or government lending agencies, they will inevitably be more "impersonal" than is customary in present farmer-country-banker relationships. "Outside" credit will probably require, and probably should receive, considerable guarantee of stability, as far as interest and principal payments are concerned. This stability feature in industry has been important in the wide acceptance of the principal of outsiders furnishing capital funds. If the principle is extended to agriculture, income stability of farmers becomes a very important factor. Incomes in agriculture need to be made more stable from two standpoints: (1) entrepreneural decisions need to be based on a "dependable" future, to insure productive efficiency on the individual farm and a resulting gain to consumers; and (2) farm incomes need to be stabilized to avoid great variations in repayment ability.

If society takes a hand in the provision of capital funds for agriculture, it will undoubtedly demand a hand in the selection of the farm entrepreneurs and in periodic examination to see how they are discharging their stewardship.[5]

Furthermore, if society takes a hand in the selection and examination of farm entrepreneurs, we will need to devote considerable study and understanding to the problem of keeping social participation democratically responsive and maintaining the greatest possible decentralization

[4] Lest this be considered too radical a departure from American practice, we should remember that society, through both public and private action, has often funished one hundred percent of the capital funds required by promotors to set up new industries. These promotors very often would not have been able to give even a fraction of the evidence of technical ability that we have customarily required for the smallest rehabilitation loan to the Farmers Home Administration!

[5] In industry, society has in some cases demanded such a hand, and in some cases not. Security and Exchange Commission regulations, public utility regulations, wages-and-hours laws, and hundreds of other welfare measures are examples of society taking a direct hand in management. An indirect hand is taken in the many cases in which some businesses are assisted, others inhibited, by tariffs, patents, and various other oligopoly positions allowed or condoned by society.

of authority and freedom of action of the individuals concerned.

Finally, society will have to safeguard itself from possible adverse consequences of its action. For example, stability in land prices and proper land-use would undoubtedly need to be achieved through legal action—otherwise this "easing" of agricultural credit could result in wild bidding up of land prices, or land might be ill-used for certain short-run purposes, when social considerations would require it to be used differently for long-run conservation ends.

Conclusion

If agricultural productive efficiency is to keep up with the demands of our expanding economy, entirely new arrangements must be devised for providing capital funds to farm operators. The point has been made here that the prospects for growth justify the creation of capital funds by the government and by banks, in much the same manner as the capital funds have been created for America's industrial growth. Both the amount of funds created and the interest rate charged for the rent of these funds are subjects for social decision. There is no valid reason for letting either of them fluctuate adversely as long as capital growth is needed in important industries.

The Problem of Allocating Capacity Cost

By RALPH K. DAVIDSON*

THE tremendous expansion in the use of electricity in the United States since the Second World War[1] brought with it a similar expansion of the productive capacity of the electric utilities[2] and a continued pressure for more expansion.[3] Seventy-nine electric utilities, owning 72% of the nation's privately-owned generating capacity, are planning to boost their capacity by 25%, according to a recent survey, at a cost of $2.1 billion.[4] Part of this pressure for expansion in capacity is a result of the American consumers' rapid postwar acceptance of air conditioning and, in some areas, the consequent drastic shifting of load curves,[5] coupled with long stretches of hot weather or "heat storms."[6] The editors of *Power Engineering*, in an annual review article entitled, "Fast Growing Load Creates Difficult Problems for Electrical Industry," declare:

"This great increase in the use of electrical energy together with the steady development in technology accounts for the tremendous progress reflected in all phases of the electrical industry. To supply this constantly increasing electrical load we need more power facilities—more and larger generating stations, longer, higher capacity transmission and switching facilities, and tremendously expanded distribution systems."[7]

After pointing out that the United States added 12.8 million kilowatts of new electric generating capacity in 1955, bringing the total to 117 million kilowatts,[8] the writers add: "By 1970, however, we will be adding new capacity at the rate of around 18 million kilowatts a year, and at that time the nation's installed generating capacity will reach 320 million kilowatts."[9]

With a total generating capacity of 117 million kilowatts, the electric utilities could have potentially produced 1,024,920,000 megawatthours during 1955. However, they actually produced only 546,838,000 megawatthours during the year,[10] a 53% utilization of capacity. In view of this under-utilization of existing capacity, a result of existing consumption patterns, one may well question the desirability of adding to capacity at the present rate. Are the present consumption patterns inevitable or are they, at least partially, the consequence of "uneconomic" rate schedules justified by incorrect allocations of capacity cost or rates based on value of service? Are the capacity costs of producing electrical energy "correctly" reflected to the users? Is it not possible, and very likely, that during certain periods users are paying less than the true economic cost for electricity and, therefore, during these periods they are consuming more

* Purdue University. The author is indebted to Professor Fritz Machlup of The Johns Hopkins University for many helpful criticisms of an earlier version of this paper. Any errors which may remain are, of course, the responsibility of the author. Part of the material on which this article is based was developed while the author held a Social Science Research Fellowship.

[1] The production and consumption of electrical energy increased from 223,177,783 megawatthours in 1946 to 546,838,000 megawatthours in 1955, a growth of 145%. "Electrical Industry Statistics," *Electrical World*, January 23, 1956, p. 156.

[2] Capacity, in megawatts, increased from 50,317 in 1946 to 114,545 in 1955, an increase of 148%. *Ibid.*, p. 160.

[3] Owen Ely, "FPC and EEI Raise Sights on Future Power Requirements and Capacity," *Public Utility Fortnightly*, December 22, 1955, pp. 1033-6.

[4] E. R. Abrams, "Electric Utilities Plan 25% Boost in Capacity," *Barron's*, February 27, 1956, p. 28.

[5] See, for example, H. L. Davis, Jr., "Room Coolers Change Load Curve," *Electrical World*, April 2, 1956, pp. 94-6; H. F. G. Main, "Room Coolers: SOS for Winter Load," *Electrical World*, March 19, 1956, pp. 136-9.

[6] C. T. Hatcher, "1955 Heat Storm Hit Brooklyn Area," *Electrical World*, April 2, 1956, pp. 90-3.

[7] *Power Engineering*, January 1956, p. 64.

[8] *Ibid.*, p. 65.

[9] *Ibid.*

[10] See footnote 1 above.

electricity than they would if they were to pay the true economic cost? In order to answer these questions it is necessary to investigate the allocation of capacity cost, a subject which has always been, and still remains,[11] highly controversial in the gas and electric industry.

In this article the author will develop an economic allocation of capacity cost for electric utilities and various alternative methods of allocating capacity cost will be analyzed and compared with an economic allocation. The special problems involved in the allocation of gas capacity cost, arising from the possibility of storing gas, will be discussed in the final section.

An Economic Allocation of Capacity Cost

During the past 30 years numerous authors have submitted various theories as to the "just" or "equitable" allocation of a utility company's capacity cost among its customers. In this article, however, the allocation of capacity cost will be analyzed from an economic point of view, leaving aside the question of "justice" and "equity." The analysis here will be centered on the question of whether or not the various methods used and proposed for the allocation of capacity cost do show the various customers the true economic cost of service.

It has been argued elsewhere that the relevant capacity cost to be allocated among a utility company's customers is the long-run marginal or "total incremental" capacity cost;[12] i.e., the expected additional cost that will be incurred if another kilowatt of capacity is provided. Included in capacity cost are interest, depreciation and maintenance of fixed equipment (the reproduction cost of the service, but not that part of the distribution system that varies with the number of customers), wages and salaries of power station operators and maintenance crews, and some managerial costs, that part of such costs that varies with the system's annual *maximum* rate of consumption in kilowatts, rather than with the *total* sales in kilowatt hours. Any allocation of average cost is essentially arbitrary, but this is not true for an allocation of marginal or incremental cost.[13] Although the authors of the various methods of allocating capacity cost were concerned with average cost, the methods they proposed could equally well be used to allocate long-run marginal or incremental capacity cost and they can be analyzed on that basis.

The basis of an economic allocation of capacity cost is the allocation to each unit of output the long-run marginal or incremental capacity cost of producing that unit of output. A customer buying a unit of output will pay, therefore, an amount sufficient to keep the resources necessary to produce that unit of output in the particular industry in the long run. At any particular time of the year the long-run marginal cost of electric energy depends on the shape of the annual load curve. Because electricity cannot be stored in appreciable quantities, the amount of capacity any electric utility company needs is determined by the maximum annual rate of consumption.

A Simple Example—Uniform Load Curve. Let us begin with a simple example in order that we may clearly see the principles of an economic allocation of capacity cost. Initially it is assumed that the expected daily load curve of a hypothetical utility company is known and that this daily load curve is exactly the same every day of the year; the utility

[11] Larry Shomaker, "Is FPC Gas Cost Allocation Equitable?" *Public Utilities Fortnightly*, November 6, 1952, pp. 680-8.

[12] Ralph K. Davidson, *Price Discrimination in Selling Gas and Electricity* (Baltimore: The Johns Hopkins Press, 1955), pp. 68-75.

[13] *Ibid.*, pp. 69, 99.

company has only one customer, with the consumption represented by the rectangle marked A in Figure 1; the company is operating in the phase of constant long-run average capacity cost; and lastly, for purposes of exposition it is assumed that the cost of the capacity required by the company is the unrealistic figure of $365 per kilowatt per year, or $1 per kilowatt per day.

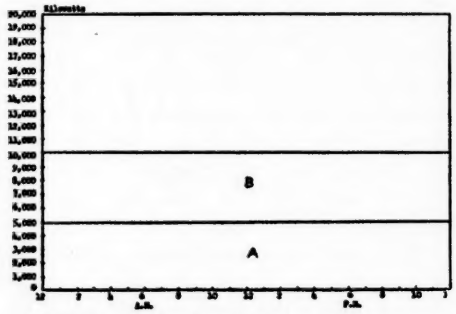


FIGURE 1—A DAILY LOAD CURVE FOR A HYPOTHETICAL UTILITY COMPANY WITH A HUNDRED PERCENT LOAD FACTOR

The single customer, A, consumes electricity steadily at the rate of 5,000 kilowatts per hour; he thus has a total daily consumption of 120,000 kilowatt hours, and his average rate of consumption is equal to his peak rate of consumption. In this case the long-run marginal capacity cost per kilowatt is spread over the twenty-four hours of a day and thus the capacity cost allocated to each kilowatt hour is equal to $1 (the assumed long-run marginal capacity cost per kilowatt per day) divided by twenty-four, or approximately 4.16c. Since A consumes at the rate of 5,000 kilowatts per hour for twenty-four hours a day, he consumes a total of 120,000 kilowatt hours each day and therefore is allocated a total capacity cost of $5,000 per day, which is equal to 4.16c times 120,000. This is the capacity cost of 5,000 kilowatts, as it should be.

Suppose another customer is added by the utility company, with the consumption marked B in Figure 1. After the company increases its capacity in order to supply the increased consumption, the marginal unit of capacity will still be used twenty-four hours a day and therefore the marginal capacity cost per kilowatt hour remains equal to approximately 4.16c. Customer B consumes electric energy at a rate equal to 5,000 kilowatts for twenty-four hours a day, a total daily consumption of 120,000 kilowatt hours; therefore, he is allocated capacity cost of $5,000 per day, i.e., 4.16c times 120,000. This is the capacity cost of 5,000 kilowatts, as it should be, since he utilizes this capacity all day. With the same rate of consumption and same total daily consumption, A pays the same as B. The addition of Customer B by the company did not change the amount of capacity cost allocated to A.

A Single-Peak Load Curve. In Figure 2 another customer is added to the previous two served by the hypothetical utility company. This customer, whose consumption is marked C in Figure 2, does not use electric energy at a constant rate, but consumes energy at the rate of 5,000 kilowatts from 6 A. M. to 10 A.M. Customers A and B still consume at the rate of 5,000 kilowatts for all twenty-four hours of the day, just as they did in Figure 1. After Customer C is added, the company has to increase its capacity from 10,000 kilowatts to 15,000 kilowatts in order to meet the increased demand for electric energy. Now, however, the marginal unit of capacity is used only four hours a day; therefore the $1 marginal capacity cost per kilowatt per day is spread over only four hours, instead of twenty-four as previously; thus, the capacity cost allocated to each unit of output consumed during the hours from 6 A.M. to 10 A.M. is equal to $1 divided by four, or 25c.

In order to supply electric energy at the maximum rate of consumption, which occurs from 6 A.M. to 10 A.M. (see Figure 2) the utility company must

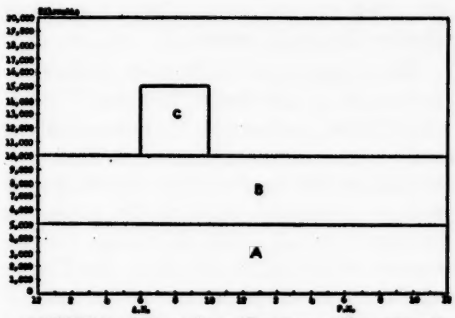


FIGURE 2—A DAILY LOAD CURVE FOR A HYPOTHETICAL UTILITY COMPANY WITH A SINGLE PEAK

maintain a plant with a total capacity of 15,000 kilowatts. During the other twenty hours of the day only part of the total capacity is utilized. Given the aggregate load curve, the long-run marginal capacity cost for energy consumed during the twenty off-peak hours is zero. There may, of course, be some increased maintenance cost, because of the use of the plant during those twenty hours, over what maintenance would be if the plant remained idle during those hours; but such increased cost is part of the marginal energy cost, and here only capacity cost is considered.

With capacity cost allocated only to energy consumed during the peak period, each of the three customers (see Figure 2) will be allocated an amount dependent upon their consumption during the peak period. Customers A, B, and C each consume 20,000 kilowatt hours during this period, i.e., at the rate of 5,000 kilowatts for four hours; therefore, they will be allocated the same amount—$5,000 each (25c times 20,000). The share of capacity cost allocated to C is equal to the increased capacity cost which his consumption brought about, and the capacity cost allocated to both A and B is unchanged. However, the calculation of the capacity cost allocated to A and B has changed; previously it was equal to 120,000 times approximately 4.16c, for both A and B, but in the situation shown in Figure 2 it is equal to 20,000 times 25c.

Effect of a Change in One Consumption Pattern. Assume now that Customer B changes his consumption pattern and no longer consumes electricity at a constant rate. In Figure 3 Customer B does not consume any electricity between 12 midnight and 4 A.M. and has doubled his rate of consumption between 3 P.M. and 7 P.M., from 5,000 to 10,000 kilowatts. The consumption patterns of Customers A and C remain as they were in Figure 2.

With these consumption patterns, the utility company no longer has a single peak but two equal peaks, one in the morning and one in the afternoon. The marginal unit of capacity is now used eight hours a day, with the capacity cost allocated to the output produced during those eight hours. It makes no difference whether the peak hours are contiguous or separated in time. In this case (Figure 3) the capacity cost allocated to each unit of output consumed during the peak period is equal to one eighth of a dollar or 12.5c per kilowatt hour.

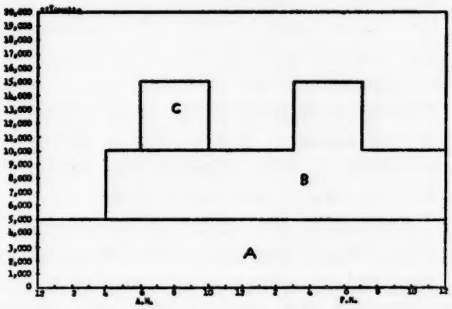


FIGURE 3—A DAILY LOAD CURVE FOR A HYPOTHETICAL UTILITY COMPANY WITH TWO EQUAL PEAKS

Notice that in an economic allocation of capacity cost all the marginal capacity cost is allocated to the output produced and consumed during the expected peak hours because the full capacity must be maintained in order to meet the peak rate of consumption during these hours. The marginal capacity cost of output during off-peak hours is zero. If output is increased during off-peak hours there will be no increase in capacity cost in the long run as long as the increase in consumption does not create a new peak, or if output is decreased during off-peak hours there will not be a reduction of capacity cost in the long run.

In the situation illustrated in Figure 3 the marginal capacity cost of a kilowatt hour of energy during the peak hours, which extend from 6 A.M. to 10 A.M. and from 3 P.M. to 7 P.M., is equal to $1 divided by eight or 12.5c. Customer A consumes at a rate equal to 5,000 kilowatts during all eight peak hours; thus he consumes a total of 40,000 kilowatt hours during the peak period, and is allocated a total capacity cost equal to $5,000, i.e., 12.5c times 40,000.

Customer B consumes at the rate of 5,000 kilowatts for four of the peak hours and at a rate of 10,000 kilowatts for the other four peak hours; he thus consumes a total of 60,000 kilowatt hours during the whole peak period. B, therefore, is allocated a total capacity cost of $7,500, i.e., 12.5c times 60,000. C consumes at the rate of 5,000 kilowatts for only four of the peak hours and thus has a total consumption during the peak period of 20,000 kilowatt hours. He is therefore allocated a total capacity cost equal to $2,500, i.e., 12.5c times 20,000.

Customer B's shift in consumption (from Figure 2 to Figure 3) did not affect the total capacity cost allocated to A, but increased the amount allocated to B and reduced the total capacity cost allocated to C. The increase in capacity cost allocated to B and the reduction to C resulted from the widening of the peak period and the consequent reduction of the amount of marginal capacity cost allocated to each kilowatt hour consumed during the peak period.

Let Customer B shift more of his consumption to the period between 3 P.M. and 7 P.M., so that the load curve for the utility company is as drawn in Figure 4. Customer A's consumption remains as it was in previous examples; he consumes at the rate of 5,000 kilowatts for all twenty-four hours in the day. In Figure 4, Customer B consumes at the rate of 3,000 kilowatts from 12 midnight to 3 P.M., at the rate of 15,000 kilowatts for the four hours from 3 P.M. to 7 P.M., and again at the rate of 3,000 kilowatts from 7 P.M. to 12 midnight. Customer C still consumes at the rate of 5,000 kilowatts between the hours of 6 A.M. and 10 A.M.

In order to be able to produce electric energy at the maximum rate of consumption during the peak period, the utility company has to increase its capacity from 15,000 kilowatts to 20,000 kilowatts and to maintain its capacity at 20,000 kilowatts. (See Figure 4) In this situation the marginal unit of capacity is used only four hours a day, from 3 P.M.

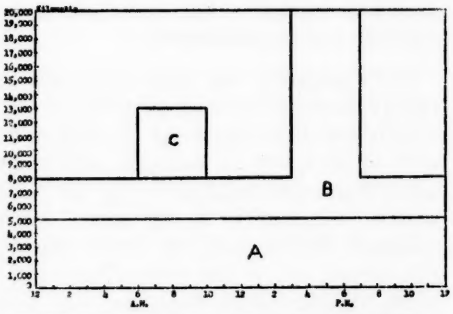


FIGURE 4—A Daily Load Curve for a Hypothetical Utility Company With a Major and a Minor Peak

to 7 P.M.; therefore, the long-run marginal capacity cost is allocated to the units of energy consumed during these four hours. The capacity cost allocated to each kilowatt hour consumed from 3 P.M. to 7 P.M. is equal to $1 divided by four or 25c. No capacity cost should be allocated to energy consumed during the other twenty hours of the day.

Customer A is allocated a total capacity cost equal to $5,000, i.e., 25c times 20,000, because he consumes electric energy at the rate of 5,000 kilowatts during all four peak hours. Customer B consumes electric energy at the rate of 15,000 kilowatts during the four peak hours and, therefore, is allocated capacity cost equal to $15,000, i.e., 25c times 60,000. All of customer C's consumption is off-peak (See Figure 4). No additional capacity is needed to satisfy C's consumption and no capacity cost would be saved if C reduced his consumption; therefore, no capacity cost is allocated to C. The amount of capacity which a utility is required to maintain to serve its customers is determined by the maximum rate of consumption during the system peak hours. Consumption during off-peak hours does not affect the amount of required capacity and therefore does not affect the capacity cost. Consequently, in an economic allocation of electric capacity cost all of the capacity cost is allocated to output consumed during the expected peak hours.

Rejection of the Peak-Responsibility Method. A variation of the peak-responsibility method of allocating capacity costs was widely used in the early period of development by electric utility companies, but later these companies rejected the method in favor of others.[14] The principal reasons given for the rejection of the peak-responsibility method were: (1) that the stability of the cost allocation was poor;[15] (2) that the annual peak shifted from one hour to another from year to year;[16] (3) that off-peak sales expanded and created a new peak or equaled the old peak;[17] and (4) that not all the customers consuming energy at the time of the system peak contributed to the peak.[18]

Capacity Cost a Function of the Load Curve. The argument that an economic allocation of capacity cost based on peak responsibility is unstable, is correct, if by unstable is meant an allocation that changes from year to year as the load curve changes; but this is not a good reason for rejecting the peak responsibility method. Capacity costs are a function of the load curve and, if the load curve changes, the change *should* be reflected in a changed allocation of capacity cost. However, part of the change in the allocation of capacity cost under the peak-responsibility method resulted from a misunderstanding and a misapplication of the method. The peak-responsibility method was not used (as it should be) to allocate capacity costs to the units of output produced during the expected peak period; instead it allocated capacity cost according to the proportion of total energy consumed by the classes of customers during a past season's peak day or

[14] Irston R. Barnes, *The Economics of Public Utility Regulation* (New York: F. S. Crofts & Co., 1947), pp. 327-8; H. W. Hills, "Demand Costs and Their Allocation," *Electrical World*, January 1927, p. 200; A. S. Knight, "Peak Responsibility as a Basis for Allocating Fixed Costs," *Electrical World*, March 1926, p. 495; L. R. Nash, *Public Utility Rate Structures*, pp. 229-30; C. Woody Thompson and Wendell R. Smith, *Public Utility Economics* (New York: McGraw Hill, 1941), pp. 390-1; Emery Troxel, *Economics of Public Utilities*, (New York: Rinehart & Co., Inc., 1947), p. 575.

[15] C. W. Barry, Prepared Direct Testimony, presented before the Pennsylvaina Public Utility Commission (C. 15763, 15733, 15742, 15758), p. 12.

[16] A. S. Knight, *Electrical World*, March 1926, p. 496; H. A. Snow, "Sharing Benefits of Diversity in Loads," *Electrical World*, March 1926, p. 404; L. R. Nash, *Public Utility Rate Structures* (New York: McGraw Hill, 1933), p. 230.

[17] H. W. Hills, "Cost of Off-Peak Business," *Electrical World*, January 1927, pp. 255-6; H. A. Snow, *Electrical World*, March 1926, p. 404.

[18] H. W. Hills, *Electrical World*, January 1927, p. 203.

peak hour. The proportions of total energy consumed by the various customers or customer classes during a past season's peak day or hour certainly need not be the same as may occur during the coming season. The correct method of determining the peak hours has been outlined elsewhere[19] and need not be repeated in detail here. Under this method the allocation of capacity cost is not determined by *ex post* day or peak hour but by the *expected* peak hours during the expected peak season, including all days which might turn out to be the peak day.

When the lighting peak was greatest the peak-responsibility method was used by utility companies to justify allocating capacity costs to lighting customers and to give special off-peak rates to industrial customers. In time the industrial load became equal to or greater than the peak lighting load. Therefore, it was argued, according to the peak-responsibility method the capacity costs should be allocated to the power peak and not the lighting peak, thus resulting in sudden large shifts in allocated costs to the different classes of service.

This criticism of the allocation of capacity cost according to peak responsibility is a result of a misunderstanding of the theoretical basis of the method. It is clear that, under the peak-responsibility method, capacity cost is allocated only to output produced and consumed during the hours when capacity is fully utilized. This implies that consumption in off-peak hours must not *exceed* the peak rate of consumption. Demand for capacity is only fully "effective" if the demander pays the full long-run marginal capacity cost. With the necessary capacity determined on this basis, it is obviously very desirable that the capacity should be used to the fullest extent possible, that is, to satisfy all demand that covers its long-run marginal energy cost, provided that *no* additional capacity is required.

Shift from Off-Peak to Peak Consumption. If the rate of consumption in a designated off-peak hour does become equal to the rate of consumption at the peak, in response to the lower off-peak rate, so that the capacity is fully utilized for an additional period of time, it does not mean that full long-run marginal capacity cost should be allocated immediately to the energy produced during that designated off-peak hour. To increase the off-peak rate to include the full long-run marginal capacity cost might well reduce the consumption of energy in that hour to a level below peak use and thus leave part of the capacity unutilized—clearly an uneconomic and inefficient use of productive resources.

If the rate of consumption in some off-peak hour becomes equal to and threatens to become greater than the rate of consumption during peak hours, when this hitherto-off-peak consumption is charged for at a price which does not include any capacity cost, the correct economic solution is to increase gradually the rate for electric energy applicable to that hour in order to restrict the rate of consumption to that which can be supplied from existing capacity. At the same time, the rate for peak energy should be reduced gradually.

In order to make this point as clear as possible, let us call the original peak hours P and designate by V the original off-peak hour in which the rate of consumption is equal to, and threatening even to exceed, the peak rate of consumption when charged for at a price that does not include any capacity cost. In the original situation, before the off-peak consumption increased, capacity cost was allocated to the units of energy pro-

[19] Davidson, *Price Discrimination in Selling Gas and Electricity* (Baltimore: The Johns Hopkins Press, 1955), pp. 184-95.

duced during the peak period, P, at a rate equal to the long-run marginal cost per kilowatt of capacity divided by the number of peak hours. When the rate of consumption during hour V increases and threatens to exceed the peak rate of consumption, V is not designated immediately a peak hour, because the demand for capacity by the customers during V is not yet really "effective" demand for it is possible that the amount of energy consumed during V is so large only because it is offered at a rate that does not include any capacity cost.

In order to keep the rate of consumption in V from exceeding the peak rate of consumption, the price for energy consumed during V is increased above the level equal to energy cost. When the price applicable to energy consumed during V, the up and coming peak hour, is increased, the revenue collected from customers consuming energy during V becomes greater than the energy cost during V. The difference between the energy cost of the electric energy consumed during V and the revenue collected from consumption during V is a credit to be applied to capacity cost as long as the capacity is fully utilized during both P and V. Capacity cost allocated to energy consumed during the old peak hours, now is equal to the total capacity cost minus the credit from consumption during V, the up and coming peak hour; that is, the capacity cost allocated to each unit of energy consumed during P is equal to the long-run marginal cost per kilowatt of capacity minus the excess of the revenue over energy cost during V, divided by the number of hours in P. If the rate of consumption during V continues to threaten to exceed the peak rate of consumption during P, the price charged for energy consumed during V should continue to be increased and the rate for energy consumed during P decreased until the point is reached when the rate per unit for energy consumed during both periods is equal. At this point, and not before, energy consumed during V is allocated full long-run marginal capacity cost. It still remains true that the long-run marginal capacity cost of energy consumed during P is equal to the long-run marginal cost of a kilowatt of capacity minus the revenue from consumption during V that is in excess of energy cost.

In other words, the capacity cost allocated to the energy produced during a particular off-peak hour, when the rate of consumption during this hour threatens to become greater than the peak rate of consumption, is increased gradually from zero to a level which, when reflected in an increased price, prevents the rate of consumption during this hour from becoming greater than the peak rate of consumption. The process of increasing the allocation of capacity cost to the former off-peak hour and reducing the allocation of capacity cost to peak hours continues until the allocation of capacity cost to output in both periods is equal per unit of output and the rate of consumption in both periods is kept equal. At this point the former off-peak hour has become a full peak hour and bears its full share of long-run marginal capacity cost. Before this point is reached, increased capacity should not be provided to satisfy the rising demand in the former off-peak hour; the rising price in the former off-peak hour keeps consumption in check while allowing the users freedom of choice of time of consumption. It was a failure to understand this gradual shifting of allocated capacity cost from peak to what had been "off-peak" consumption but was now threatening to rise above the peak rate of consumption, plus an incorrect determination of what consti-

tuted the peak, that led to the plea to reject the peak-responsibility method of allocating capacity cost.[20]

The argument that some customers who consumed energy at the time of the system peak do not contribute to that peak[21] is completely false. Hills argued that a customer with a 100% load factor "does not cause any peak for he has none, and any station peak that occurs is due entirely to influences over which he has no control."[22] If such a customer, or any customer who uses energy during the system peak, were to transfer his consumption from the system peak to a different point of time, the system peak would be smaller and therefore less capacity would be needed by the company so that capacity cost would be less. All customers who consume energy during the system peak are co-responsible for the peak, according to their consumption at that time.

Alternative Proposals

In this section we propose to analyze these various methods and then to compare them with the economic allocation we proposed in the first section.

W. J. Greene's Consumption and "Demand" Method. The first of these alternatives to be discussed is W. J. Greene's consumption and "demand" method.[23] Greene allocates part of the capacity cost to the various customers according to their individual maximum rates of consumption, whenever they occur, and the rest of the capacity cost to the customers according to their individual total consumptions. Very serious flaws are present in this scheme for not only do off-peak users subsidize the peak users but some customers are able to make a shift in consumption which will increase the system capacity cost while reducing the capacity cost that is allocated to themselves.

For his consumption and "demand" method, Greene sets up two equations:

Kx plus Dy equals C and

8,760x plus y equals C divided by P

The symbols in the equations mean:

C equals the total annual capacity cost incurred by the system, part of which Greene allocates according to the maximum rates of consumption of the individual customers and the remainder of the capacity cost according to total annual consumption of the individual customers.

K equals the number of kilowatt hours used by all the customers in a year.

x equals the share per kilowatt hour of the portion of the capacity cost which Greene allocates to the total number of kilowatt hours supplied by the system during a year. Therefore, Kx is equal to the portion of total capacity cost that Greene allocates according to total consumption.

D equals the sum of the individual customer's maximum "demands," irrespective of the time they occur.

y equals the share per kilowatt of the portion of capacity cost that Greene allocates according to

[20] W. Arthur Lewis presents the core of the correct solution in a very abbreviated form in his article, "The Two-Part Tariff," *Economica*, August 1941, pp. 250-1, where he uses the example of cotton seed and lint as an illustration of the problem of allocating cost between peak and off-peak periods. However, he argues that "if a mere lowering of price in slack periods stimulates demand sufficiently to keep equipment fully occupied at a price greater than zero, no exact allocation of costs is possible between peak and slack." We have argued above that in such a situation the marginal cost of P was equal to the long-run marginal cost of a kilowatt of capacity minus the net marginal revenue of V. After the above analysis was worked out, it was drawn to our attention that a substantially similar solution was given by Lewis in the expanded version of his article on the two-part tariff which appears in *Overhead Costs* (London: Allen & Unwin, 1949), pp. 48-9. Referring to his cotton seed-lint example, he writes, "There are three stages to the process: a stage where the whole cost is attributable to lint, a stage where the whole cost is attributable to seed, and an intermediate stage where cost allocation depends on demand."

[21] H. W. Hills, *Electrical World*, January 1927, p. 203.

[22] *Ibid.*, p. 203.

[23] W. J. Greene, "Determining Demand Charge," *Electrical World*, November 1925, p. 947.

the individual customer's maximum "demand," irrespective of the time of occurrence. Dy, therefore, is equal to the portion of total capacity that Greene allocates according to customers' maximum "demands."

P equals the maximum coincident "demand" or peak responsibility of all customers on the sources of supply, i.e. it is equal to the system peak.

8,760 equals 24 x 365, or the number of kilowatt hours used in a year by a 1-kilowatt load operated at 100 per cent power factor and at 100 per cent load factor.[24]

The relevant data for any system are put into the two equations, which are then solved for x and y. Using Greene's method, the capacity cost allocated to a customer is equal to the sum of the individual's maximum rate of consumption multiplied by y and his total consumption in kilowatt hours multiplied by x.

Consider, for example, the daily load curves shown in Figure 5,[25] in which all

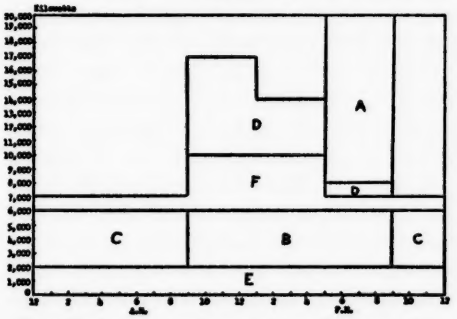


FIGURE 5—Allocation of Capacity Cost: Case I

the customers have exactly the same total daily consumption but different load factors. For simplicity it is assumed that each customer has the same consumption pattern every day of the year, so that daily consumptions can be used instead of annual consumptions and 24 can be substituted for 8,760 in Greene's equation. Using Greene's equations thus modified, with the data depicted in Figure 5, the capacity responsibility, in kilowatts, for the company's required capacity of 19,000 kilowatts is allocated among the company's customers as follows: A, 5,183; B, 2,636; C, 2,636; D, 3,909; E, 2,000; and F, 2,636, which equals a total of 19,000 kilowatts. The capacity responsibility times the long-run marginal cost per kilowatt of capacity per day is equal to the allocated capacity cost. However, in the discussion which follows, we shall use capacity responsibility to compare the different methods of cost allocation.

Let us assume now that Customer D changes his pattern of consumption but does not change his total consumption, by shifting 4,000 kilowatt hours of his consumption from the period 9 A.M.-1 P.M. to the period 5 P.M.-9 P.M. As a result of this shift in D's consumption the system peak rate of consumption has increased 1,000 kilowatts, raising the required capacity by 1,000 kilowatts, from 19,000 to 20,000 kilowatts, as shown in Figure 6. Using the data depicted in Figure 6 and Greene's modified equations, the total responsibility for the 20,000 kilowatts of capacity is allocated among the company's customers as fol-

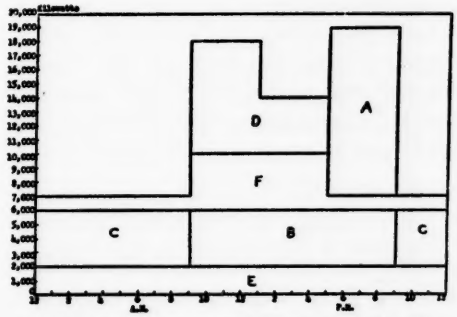


FIGURE 6—Allocation of Capacity Cost: Case II

[24] *Ibid.*, p. 947.

[25] Figure 5 is a copy of a figure first used by John Oram and H. H. Robison, "The Complete Peak Method," *Electrical World*, August 1928, p. 359, and then again by R. E. Purucker in a paper presented at the 29th Annual Convention of Utility Commission Engineers, June 1951, "Electric Cost Allocations," p. 16.

TABLE I—A Comparison of the Capacity Responsibility Allocated Under the Various Methods (Refer to Figures 5 and 6) *

Method	A		B		C		D		E		F	
	I	II	I	II	I	II	I	II	I	II	I	II
Greene	5,183	5,809	2,636	2,762	2,636	2,762	3,909	3,905	2,000	2,000	2,636	2,762
Hills	7,833	8,667	3,167	3,333	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000
Oram and Robison	5,055	6,306	3,366	3,565	2,000	2,000	3,822	3,425	2,000	2,000	2,757	2,704
Purucker	4,316	5,100	3,555	3,734	2,000	2,000	3,375	3,341	2,778	2,867	2,976	2,958
Nissel	5,083	6,194	3,861	4,046	1,000	1,000	4,111	3,852	2,000	2,000	2,944	2,908
Non-Coincident Demand	6,705	7,273	2,235	2,424	2,235	2,424	4,470	4,243	1,120	1,212	2,235	2,424
Nordin	4,105	4,800	3,590	3,777	2,000	2,000	3,445	3,459	2,795	2,888	3,065	3,076
Economic Peak Responsibility	12,000	12,000	4,000	4,000	0	0	0	1,000	2,000	2,000	1,000	1,000

* I refers to Figure 5 and II refers to Figure 6.

lows: A, 5,809; B, 2,762; C, 2,762; D, 3,905; E, 2,000; and F, 2,762. As a result of the shift, the capacity responsibility, and, consequently, the capacity cost allocated to D has decreased, although his shift in consumption has increased the total capacity cost for the system. Under Greene's method, any customer whose rate of consumption at the time of the system peak is less than his maximum rate of consumption can make such a move as D's without increasing his capacity responsibility.

Greene also developed another method which he termed the "used and unused capacity" method but, as he demonstrated, the two methods give identical results; therefore the second will not be discussed.[26]

The "Phantom Customer" Method. H. W. Hills presented the "phantom customer method" in two articles in the *Electrical World* in 1927.[27] The areas marked A, B, C, D, E, and F, in Figure 5, represent the actual consumption of a utility company's customers. In Hills' method, the rest of the area, which is the difference between actual consumption and what consumption would be with a 100% load factor for the system, represents the imaginary consumption of a "Phantom customer." The company's total capacity cost is allocated equally over all the consumption, including the imaginary consumption of the phantom, i.e., as if the company operated with a 100 % load factor. Then Hills allocates the phantom's share of the company's capacity cost to the peak load customers, in proportion to their contribution to the peak, according to his definition of peak responsibility. In Hills' scheme each customer contributes to the system peak the amount of the excess of his rate of

[26] W. E. Greene, "Allocating Capacity Costs, *Electrical World*, May 1926, pp. 1190-1.

[27] H. W. Hills, "Demand Costs and Their Allocation," *Electrical World*, January 1927, pp. 198-203; H. W. Hills, "Proposed Allocation of Demand Costs," *Electrical World*, January 1927, pp. 249-52.

consumption at the time of the system peak over his average rate of consumption.[28]

Using the data underlying Figure 5, the capacity responsibility in kilowatts is allocated to the customers, according to Hills' phantom customer theory, as follows: A, 7,833; B, 3,167; C, 2,000; D, 2,000; E, 2,000; and F, 2,000. Again let us assume that customer D, in Figure 5, changes his consumption to that of D, in Figure 6. As a result the peak rate of consumption for the system has been increased, necessitating an increase in capacity if the company is to be able to meet the peak rate of consumption.

Under Hills' method the allocation of capacity responsibility, after the shift in D's consumption, is as follows: A, 8,667; B, 3,333; C, 2,000; D, 2,000; E, 2,000; and F, 2,000. The capacity cost allocated to D has not been increased although his action caused an increase in capacity required by the utility company to meet the peak rate of consumption. Any customer whose rate of consumption at the time of the system peak is less than his average rate of consumption is able to make such a shift in consumption without increasing his own allocation of capacity cost, under Hills' method.

The "Complete Peak" Method. Under the "complete system peak" method advocated by John Oram and H. H. Robison, "complete system peak" includes all the periods of time in which the system load is greater than the average load on the system.[29] The authors begin by saying that whenever the system load exceeds the average load this must be because some customers at some times have a rate of consumption in excess of their average rate of consumption. Therefore, they argue, the capacity cost of the company's plant during the time of idleness should be allocated to those customers using energy during the "complete peak" period in proportion "to their respective excess demands over their average demands during each interval of the peak."[30]

The Oram-Robision method allocates the total capacity cost to each hour on the assumption of 100% load factor for the system, as in the Hills' method, but the capacity cost of the unused hours is allocated differently. To illustrate the method, the same customers, with the same load curves that are shown in Figure 5 are given in Figure 7, plus some additional material.[31] The areas marked A, B, C, D, E, and F, represent the consumption of the six customers, exactly the same as in Figure 5. The shaded area represents the unused capacity of the company. The unused kilowatt hours are allocated to the hours of the "complete peak," which are the hours during which the system load is greater than the average system load, 9 A.M. to 9 P.M. in Figure 7, as follows: Area 1,

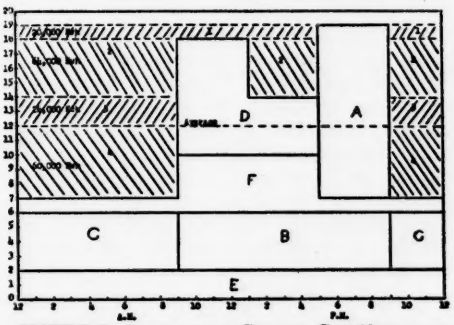


FIGURE 7—ALLOCATION OF CAPACITY COST UNDER THE "COMPLETE PEAK" METHOD

which is equal to 20,000 kilowatt hours or 1,000 kilowatts of capacity used for 20 hours, is added to the consumption taking place from 5 P.M. to 9 P.M.

[28] *Ibid.*, p. 250.

[29] John Oram and H. H. Robison, "The Complete Peak Method," *Electrical World*, August 1928, pp. 359-61.

[30] *Ibid.*, p. 359.

[31] Figure 7 is a copy of Figure 2 in the article by Oram and Robison, *ibid.*, p. 360.

Area number 2, a total of 64,000 kilowatt hours, is equally divided among the hours from 9 A.M. to 1 P.M. and 5 P.M. to 9 P.M. Area number 3, 24,000 kilowatt hours in total, is allocated equally to the hours from 9 A.M. to 9 P.M. Area number 4 is treated differently. All of Area number 4, in Figure 7, lies below the average load of the company and is therefore equal to the number of kilowatt hours in the system load which lie above the system average. Oram and Robison therefore allocate Area number 4 to the consumption which lies above the average on a kilowatt hour per kilowatt hour basis, i.e., each kilowatt hour which lies above the system average is allocated the capacity cost originally allocated to a kilowatt hour in Area number 4.

The unused kilowatt hours, after being allocated to the hours of the "complete peak," are than allocated to the different customers who consume energy during the "complete peak" period. The allocation of the unused kilowatt hours among the customers is based on the following argument:

> "The reason for the load exceeding the average, in the curve considered, is because certain systems have a demand in excess of their average during the 9 A.M. to 9 P.M. hours. Therefore, the systems that use more than their average demand during this period are responsible for the peak and their degree of responsibility is in direct proportion to their respective excess demands over their average demands for each period of the peak. In other words, any class that exceeds its average demand when the total system load is in excess of the average is a peak-producing class and its costs should be allocated according to these simultaneous excesses."[32]

The actual kilowatt hour consumption of each customer, plus his allocation of the unused kilowatt hours, is then divided by 24 to find the capacity responsibility of each customer. It was assumed that the daily load curves are constant over the year, so 24 can be used instead of 8,760. Using Oram and Robison's "Complete Peak" method, the capacity responsibility in kilowatts, is allocated to each customer, in Figure 5, as follows: A, 5,055; B, 3,366; C, 2,000; D, 3,822; E, 2,000; and F, 2,757.[33]

Assume now that D shifts 4,000 kilowatt hours of his consumption from the period 9 A.M.-1 P.M. to the period 5 P.M.-9 P.M. as in Figure 6. This shift increases the maximum rate of consumption on the company by 1,000 kilowatts, necessitating an increase in capacity of that amount in order to satisfy peak requirements. Allocating the capacity responsibility in Figure 6 according to Oram and Robison's method, the following capacity allocations in kilowatts result: A, 6,306; B, 3,565; C, 2,000; D, 3,425; E, 2,000; and F, 2,704. It is to be noted that by his action D reduced his allocated capacity responsibility although his action increased the capacity requirements for the company. Such a shift in consumption, which is detrimental to the company but beneficial to the customer, is possible under the Oram and Robison method whenever a customer's peak rate of consumption occurs during the "complete peak" but not during the true peak rate of consumption for the system as a whole. By shifting part of his consumption from his own peak period to the peak period of the company a customer can reduce his bill by reducing the excess of his maximum rate of consumption over his average rate of consumption and yet increase the total rate of consumption during the system peak hours.

A "Modified Complete Peak" Method. R. E. Purucker, of the staff of the Public Service Commission of Wisconsin,

[32] *Ibid.*, p. 360.

[33] *Ibid.*, p. 360.

suggested modification of Oram and Robison's "complete peak" method in his paper on electric cost allocations presented before the Conference of utility Commission Engineers in June, 1951.[34] Purucker follows Oram and Robison exactly in his determination of the complete peak period and the assignment of the unused capacity to the hours of the complete peak. He differs from them, however, in his allocation of the unused capacity to the various customers. He proposes that the unused capacity assigned to the complete peak hours be apportioned to the customers on the basis of used capacity, or the consumption of each customer or group of customers during each period of time in the complete peak.[35] According to Purucker's modified complete peak method, the allocation of the capacity responsibility in the case shown in Figure 5 (which Purucker also uses), is as follows: A, 4,316; B, 3,555; C, 2,000; D, 3,375; E, 2,778; and F, 2,976.[36]

Purucker's method will now be used to allocate capacity responsibility for the data shown in Figure 6, which differs from Figure 5 in only one respect: D has shifted 4,000 kilowatt hours of his consumption from the period 9 A.M.-1 P.M. to the period 5 P.M.-9 P.M., thereby increasing the maximum rate of consumption for the company as a whole by 1,000 kilowatts, necessitating an increase in capacity. The following allocation of capacity responsibility, in kilowatts, results: A, 5,100; B, 3,734; C, 2,000; D, 3,341; E, 2,867; and F, 2,958. As in the other methods discussed, the capacity responsibility allocated to D has been reduced by his shift in consumption, although his move increased the capacity cost for the system as a whole.

The "Capacity Credit" Method. Another modification of Oram and Robison's complete peak method has been proposed by Hans E. Nissel.[37] Although Nissel claims his method is designed especially for gas utilities with interruptible customers, his method is applicable to either gas or electric utilities. The specific problem of allocation of capacity cost to interruptible-gas customers will be considered at a later point. Nissel also points out that his system uses the annual load curve rather than the load curve for the peak day of the year, but in explaining his method it will be assumed that the daily load curves in Figures 5 and 6 are the same every day in the year.

Under the "capacity credit" method, the unused capacity is allocated to the "complete peak" hours, as in Oram and Robison's scheme. Nissel uses a different system to allocate the unused capacity, but the results are exactly the same. The novel feature of Nissel's "capacity credit" method is that any customer (or group of customers) receives a capacity credit if his rate of consumption is below his average rate of consumption during any time interval of the complete peak period, and an amount equal to the credit is charged to the other customers using energy during that period. For example, in Figure 5, F's rate of consumption from 5 P.M. to 9 P.M. is less than his average rate of consumption; therefore he receives a credit to the extent of his deficiency, in this case 167 kilowatts. After the unused capacity has been assigned to the hours making up the complete peak and all capacity credits

[34] R. E. Purucker, "Electric Cost Allocation," a paper presented at the 29th Annual Conference of Utility Commission Engineers at Dearborn, Michigan, June 27, 28 and 29, 1951.

[35] *Ibid.*, p. 15.

[36] *Ibid.*, p. 18.

[37] Hans E. Nissel, "Technique and Application of New Capacity Credit Method," *American Gas Journal*, November 1953, pp. 16-17, 30-1. This method was used in an analysis of total cost presented by The Peoples Gas Light and Coke Company before the Illinois Commission (Exhibit No. 34, Docket No. 38244).

have been given, the unused capacity assigned to the complete peak hours is allocated to the various customers consuming energy during the complete peak "in proportion to the excess of their demands over their individual averages."[38]

The capacity credit method gives the following capacity responsibility, in kilowatts, for the utility customers in Figure 5: A, 5,083; B, 3,861; C, 1,000; D, 4,111; E, 2,000; and F, 2,944. Using the capacity credit method to allocate capacity responsibility to the customers in Figure 6—where D has switched some of his consumption, increasing the system's capacity cost—the following allocation in kilowatts results: A, 6,194; B, 4,046; C, 1,000; D, 3,852; E, 2,000; and F, 2,908.

Under the capacity credit method there are two forces at work to reduce the capacity allocated to D as a result of his shift of consumption. First, his move reduces the excess of his rate of consumption during the 9 A.M-1 P.M. period over his average rate of consumption, thus reducing his share of the capacity assigned to that time period. Secondly, although D's shift of consumption reduced his capacity credit during the 5 P.M.-9 P.M. period, this is more than offset by the shift in capacity responsibility from the 9 A.M.-1 P.M. period, in which D shares, to the 5 P.M.-9 P.M. period, during which D receives a capacity credit. Thus, again, under the capacity credit method a customer receives a bonus, as it were, as a result of a move that increases the capacity cost for the system as a whole.

The "Non-Coincident Demand" Method. The "non-coincident demand" method is now probably more commonly used in cost-of-service studies than any other method of allocating capacity cost.[39] Under this method of allocating capacity cost the maximum rates of consumption of all customers or classes are added together irrespective of the time of occurrence to find the aggregate "non-coincident demand." Capacity cost is then allocated to each class according to the ratio which the maximum rate of consumption, at whatever time of day, of that class bears to the aggregate "non-coincident demand."

Using the "non-coincident demand" method of allocating capacity responsibility for the case illustrated in Figure 5, the following allocations, in kilowatts, result: A, 6,705; B, 2,235; C, 2,235; D, 4,470; E, 1,120; and F, 2,235. Again, assume that D shifts some of his consumption, so that the system load curve and individual consumptions are those given in Figure 6. The "non-coincident demand" method gives the following capacity responsibility allocation: A, 7,273; B, 2,424; C, 2,424; D, 4,243; E, 1,212; and F, 2,424. Although D's action resulted in an increase in the capacity required by the utility company in order to meet peak "demand," the capacity responsibility allocated to D is again reduced from Figure 5 to Figure 6.

Under the "non-coincident demand" method such a move is possible whenever an individual's maximum rate of consumption occurs at a time other than the system peak. The individual customer can reduce his capacity responsibility by reducing his maximum rate of consump-

[38] *Ibid.*, p. 17.

[39] C. W. Bary, Prepared Direct Testimony before the Pennsylvania Public Utility Commission, Docket C 15763. The "non-coincident demand" method was used by Bary in a cost-of-service study for the Philadelphia Electric Company in the above case. This method was also used in the cost-of-service study presented in Gas Rate Case No. D-2948, The Consumers Power Company before the Michigan Commission and in the cost-of-service studies prepared by A. F. Rehberg, for the Consolidated Edison Company of New York, and J. D. McKechnie, of the New York Public Service Commission staff, in connection with the Consolidated's rate case—Case 12455—before the New York Public Service Commission.

tion and transferring some of his consumption to the system's peak period, an action which is detrimental to the system as a whole, but beneficial to himself. Another flaw of the "non-coincident demand" method is seen with reference to customer E, in Figures 5 and 6. Although E uses a capacity of 2,000 kilowatts continuously, he is allocated responsibility for only 1,120 kilowatts in Figure 5 and 1,212 kilowatts in Figure 6.

Nordin's Method of Allocating Capacity Cost. A method of allocating capacity cost that is specifically designed to induce customers to shift their consumption from the utility company's peak period has been developed by J. A. Nordin.[40] Under Nordin's method capacity responsibility for each increment of capacity is allocated to the hours during which that increment of capacity is used. The allocations to each hour are summed to find the total hourly allocation, which is then divided by the number of increments of capacity in use during each hour to find the average hourly capacity responsibility—which varies from hour to hour depending upon the shape of the load curve. The hours of consumption of each customer (or class) are then multiplied by the average hourly capacity responsibility for each particular hour to find the allocated capacity responsibility for the customer (or class).

For example, in Figure 5, the responsibility for the first 7,000 kilowatts of capacity (beginning from the x axis) is allocated evenly over all twenty-four hours. The responsibility for the next 7,000 kilowatts of capacity is allocated equally among the twelve hours from 9 A.M. to 9 P.M. The next 4,000 kilowatts are allocated to the hours between 9 A.M. and 1 P.M. and 5 P.M. and 9 P.M., and the last 1,000 kilowatts of capacity are allocated solely to the hours from 5 P.M. to 9 P.M. Average hourly capacity responsibility for the hours from 5 P.M. to 9 P.M., for example, is thus 13/152nds of a kilowatt. Customer A, who has a rate of consumption equal to 12,000 kilowatts for these four hours, is allocated capacity responsibility for 4,105 kilowatts. Nordin's method yields the following capacity responsibilities in kilowatts, for the customers in Figure 5: A, 4,105; B, 3,590; C, 2,000; D, 3,445; E, 2,795; and F, 3,065.

Using the data shown in Figure 6 and allocating the capacity responsibility according to Nordin's method, the following allocation of capacity responsibility, in kilowatts is given: A, 4,800; B, 3,777; C, 2,000; D, 3,459; E, 2,888; and F, 3,076. Nordin is partially correct when he says that under his system customers will have "the least incentive to shift to the peak,"[41] because his system is designed to increase a customer's capacity cost allocation when he makes a shift in consumption that adds to total system capacity cost. Similarly, Nordin's system will reduce the capacity responsibility allocated to a customer if the customer will shift some consumption from the peak to off-peak period. Although Nordin's method is a step in the right direction, it is far from an economic allocation of capacity cost. In the cases illustrated in Figures 5 and 6, for example, the shift in consumption by customer D from a system off-peak to a system peak period necessitated an increase in the company's capacity of 1,000 kilowatts in order to meet the new peak "demand," but under Nordin's method D's capacity responsibility was increased by only 14 kilowatts, not by the full 1,000.

Capacity Allocations under the Various Methods Compared. The allocations of capacity responsibility under the various

[40] J. A. Nordin, "Allocating Demand Costs," *The Journal of Land & Public Utility Economics*, May 1946, pp. 163-170.

[41] *Ibid.*, p. 167.

methods for the data underlying Figures 5 and 6 are brought together in Table I in order to compare the results. The most striking fact to be observed is the great extent to which the peak users are subsidized by the off-peak customers under all the alternative methods of capacity cost allocation. In Case I (Figure 5) five of the seven methods, alternative to the economic allocation of peak responsibility, allocate to Customer A less than half of his full peak responsibility. Nordin's method allocates the least capacity responsibility to A, slightly more than ⅓ of his full peak responsibility, and Hills' method, which allocates to A more of the capacity responsibility than any of the other methods, still allocates to him less than ⅔ of his full responsibility.

The extent to which off-peak customers subsidize peak customers under all the alternative methods is clearly seen in their treatment of customer C, a purely off-peak customer under all definitions of the peak. Four of the methods allocate to C his so-called "used" capacity, while two of the methods charge C a higher amount; in actual fact C has absolutely no peak responsibility. Customer D, another off-peak customer in Figure 5, is charged most under the "non-coincident demand" method and least under the Hills method. Four of the methods allocate the correct capacity responsibility to E, a 100% load customer; two of the methods allocate almost 50% more to E, and the "non-coincident demand" method allocates to him almost 50% less than his peak responsibility, although he consumes at the same rate twenty-four hours a day and therefore "uses" the same amount of capacity all the time.

A very damaging criticism of all the alternative methods is suggested when the capacity responsibility allocations of Figure 5 and Figure 6 are compared.

The method advocated by Nordin is the only one that increases the capacity responsibility allocated to D in going from Case I (Figure 5) to Case II (Figure 6). Of the other six methods, Hills' leaves D's capacity allocation constant, while all the other methods reduce the capacity responsibility allocated to D; the greatest decrease is under Oram and Robison's method. Although Nordin's method does increase the capacity responsibility allocated to D when he shifts some of his consumption to the peak period, the increase may be relatively small; in our example responsibility for only 1.4% of the increased capacity necessitated by D's action was allocated to D.

If the demand for electric energy has any responsiveness to price at all, the use of any of the alternative methods of allocating capacity responsibility and, accordingly, of capacity cost, in determining electric rates will bring about a decreased use of existing facilities, an increased investment in new facilities, and a more inefficient use of the new plant, than would exist under an economic allocation of capacity cost, with rates based on cost of service. Wherever rate schedules do not reflect the variation of cost between peak and off-peak service the users of electricity during peak periods will be subsidized by the users of electricity during off-peak periods whether or not the actual rate schedules are based on an incorrectly calculated cost of service or on value of service, i.e., economic price discrimination. It appears very likely that part of the tremendous expansion of investment in electric capacity has been induced by below-cost rates for electrical energy consumed during peak hours. This follows directly from the existing subsidization of consumption of electricity during peak hours by those who consume during off-peak hours. In the section to

follow, problems involved in an allocation of capacity cost for a gas utility company will be discussed with particular attention to the special problems created by the economic storage of gas.

Complications Arising from Storage of Gas

The theoretical explanation of an economic allocation of capacity cost for gas utilities would be exactly the same as that given for electric utilities in the first section of this article, if gas could not be stored. The economically correct method of allocating gas capacity cost would be to allocate the marginal or incremental capacity cost to the units of output produced during the expected peak period. However, gas can be stored and has been on a markedly increasing scale since World War II;[42] and this introduces a new element into the theoretical analysis of an economic allocation of gas capacity cost and one which, up to now, has not been recognized.

Since the gas distribution system must be large enough to satisfy peak-hour rates of consumption, the economically correct allocation of distribution capacity cost follows the peak-responsibility method as explained in the first section, and need not be repeated here. However, the fact that gas can be stored complicates matters for the allocation of gas production capacity cost, including in gas "production capacity" cost the buying of natural gas from a pipeline company if the company distributes natural gas instead of locally maufactured gas.

In order to illustrate the principles of the allocation of gas capacity cost, it is necessary to divide the capacity cost into three components, production capacity, storage capacity, and distribution capacity cost. As has been indicated, the allocation of distribution capacity cost is exactly the same as the allocation of electric capacity cost developed in the first section and will not be considered further. The remaining part of this section will deal with the allocation of production and storage capacity cost.

Assume that the utility company represented in Figure 8 is a gas utility company and the axis of ordinates measures capacity in cubic feet of gas. It is assumed that the daily load curve shown in Figure 8 is constant for every day of the year; each customer consumes 48,000 cubic feet of gas per day so the total daily consumption of all customers is 288,000 cubic feet of gas; the utility company distributes natural gas and buys its requirements from a pipeline company under a Hopkinson-type rate schedule. Since gas can be stored, the company's maximum rate of taking from the pipeline need not be more than 12,000 cubic feet per hour, i.e., the total daily requirement of 288,000 cubic feet divided by twenty-four hours. Under these circumstances the company would need storage capacity for 60,000 cubic feet. From 9 P.M. (in Figure 8) until 9 A.M. gas will flow into storage to be used later to meet the consumption requirements from 9 A.M. to 9 P.M. Under this arrangement the company maintains a 100% load factor for its taking from the pipeline. In this case all the production capacity cost, including the payment by the local gas distributing company to the pipeline company for

[42] Harold W. Springborn, "Underground Storage Exceeds Two Trillion," *Gas Age*, May 3, 1956, pp. 29-36. Mr. Springborn writes that "total estimated ultimate gas in storage passed the two trillion cubic foot mark this year." p. 29. "With 214 storage fields located in 20 states, this phase of the total gas industry operation has required more than $270,000,000 in construction expenditures during the last six years, and will demand another $113,000,000 in the three years 1956-1958." p. 29.

Gene T. Kinney, "Underground Gas Storage Still Rising," *The Oil and Gas Journal*, April 23, 1956, pp. 68-71. "Importance of storage . . . . although 41 years old, didn't become an important factor in the natural-gas industry until after World War II. Capacity has since multiplied more than tenfold. It has grown to the extent that storage will supply as much as half of the space-heating requirements for some gas companies during the coldest days of the winter." p. 68.

natural gas, is allocated to all twenty-four hours of the day, since gas not sold to a customer in an off-peak hour will go into storage to be sold during the peak period and thus reduce the delivery of natural gas from the pipeline during the peak period.

Storage capacity cost, in Figure 8, is chargeable to the gas consumed during

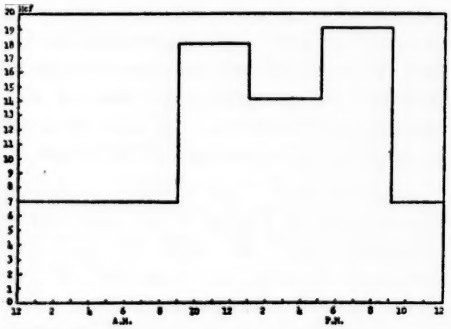


FIGURE 8—Allocation of Gas Capacity Cost

the hours in which the company's load curve is in excess of its average value, and not only the consumption during the hours of the true peak period, because a reduction in consumption during any of the hours in which the load curve is in excess of its average value would mean a reduction in the required storage capacity.

If the company's load curve has more than one peak, and the company has to provide a greater storage capacity than is necessary to meet only the consumption during the largest peak, storage capacity cost is allocated to consumption taking place during all the hours for which the load curve is in excess of its average value. But if, although the load curve has more than one peak, the company does not have to provide additional storage capacity beyond that required in order to satisfy requirements during the hours of the largest peak, storage capacity cost is allocated only to the output consumed during the hours of the largest peak.

Let us now assume that it is cheaper for the gas utility company, represented in Figure 8, to maintain storage capacity for 36,000 cubic feet of gas rather than 60,000 cubic feet. With this amount of storage capacity the company will need to have the gas holders full of gas at the start of the morning peak (9 A.M.), shown in Figure 8, and draw on the pipeline at the rate of 14,000 cubic feet per hour from 9 A.M. to 9 P.M. By increasing its maximum rate of delivery of gas from the pipeline from 12,000 cubic feet per hour to 14,000 cubic feet per hour, while still taking the same total quantity of gas during the day, the company is able to reduce its storage requirements 24,000 cubic feet, i.e., storage requirements are reduced 12 cubic feet for every 1 cubic foot per hour increase in the rate of delivery from the pipeline. When the rate of delivery of gas from the pipeline is at the level of 14,000 cubic feet, in Figure 8, an additional increase of 1 cubic foot in the rate of delivery will only reduce storage requirements by 8 cubic feet. For this latter increase in the company's maximum rate of consumption, it is assumed that the increase in "demand" charges under the Hopkinson rate, assumed to be used by the pipeline company, will add more to total costs than will be saved by the reduction in storage capacity, but the reverse is assumed to be true for the increase in the maximum rate of delivery from 12,000 to 14,000 cubic feet per hour.

From 9 A.M. to 1 P.M. the company will send out 16,000 cubic feet of gas from storage, and from 5 P.M. to 9 P.M. it will send out 20,000 cubic feet of gas from storage, in addition to the gas received from the pipeline at the rate of 14,000 cubic feet per hour. To be ready for the following day's business, the com-

pany will need to refill the gas holders during the off-peak hours from 9 P.M. to 9. A.M.

In the present example the production capacity cost is allocated to the hours from 9 A.M. to 9 P.M. and the storage capacity cost is allocated to the hours in the periods 9 A.M.-1 P.M. and 5 P.M.-9 P.M. No capacity cost is allocated to the hours from 9 P.M. to 9 A.M., since any reduction in consumption during those hours will reduce neither the required production capacity cost nor the required storage capacity.

Capacity Cost Allocation to Interruptible Gas Customers. Interruptible gas customers are those customers who buy gas on the understanding that their consumption may be cut off on very short notice when the production capacity is needed in order to meet the requirements from the so-called "firm" customers.[43] Therefore, it is argued, since interruptible customers do not contribute to the peak rate of consumption no part of capacity cost should be allocated to such interruptible customers.[44] This argument is only correct if the gas utility company is unable to store gas; but where gas can be stored in appreciable quantitites the argument is invalid with respect to production and storage capacity costs and valid only for distribution capacity cost.

Consider the first gas example used above, shown in Figure 8, with capacity in cubic feet of gas measured on the ordinate axis. In this case the company buys natural gas from a pipeline at the rate of 12,000 cubic feet per hour, or 288,000 cubic feet per day, with a 100% load factor, and has storage capacity for 60,000 cubic feet. (This example is very similar analytically to a case cited below.[46]) Under these circumstances, if the company were to offer interruptible gas service to any customer, it would have to provide additional production capacity, the same as for any other of its customers, because the gas the company receives from the pipeline each day is just sufficient to meet the daily consumption of its customers. Any increase in sales during any hour of the day could only be met by an increase in the rate of delivery from the pipeline to the gas utility company.

Alternatively, the gas company could purchase gas from the pipeline company at a rate equal to the maximum rate of consumption of its regular firm customers and dispose of the gas not required by the firm customers during some hours by selling this "excess" gas on an interruptible basis. And then the company could argue that the production capacity was required in order to meet the peak rate of consumption of its firm customers and therefore all capacity cost should be

[43] "The principal distinction between general firm customers and interruptible and off-peak users is the character of the service in that the off-peak customers are firm, but only for a specified portion of the year, and the interruptible customer uses gas during all of the year but only at times when it is available because not taken by the firm customers." Report of the Subcommittee on Interruptible, Off-peak and Seasonal Gas Rates, *Proceedings*, American Gas Association (1952), p. 70.

[44] Off-peak and interruptible loads are controlled in such a way that they do not cause an increase in maximum demand and, consequently, in capacity cost. They have, therefore, no definite capacity responsibility for jointly used plant." *Report of Rate Committee* (New York: American Gas Association, 1953), p. 59. This is a summary review of a cost-of-service study prepared by The Peoples Gas Light and Coke Company.

[46] ". . . . Although the input deliveries into the California section at Blythe are maintained at a substantially uniform rate at approximately 305 M³cf/day, it is possible to vary terminal deliveries widely from hour to hour in accordance with the requirements of load conditions in the Los Angeles area. Deliveries from the line during peak load hours often exceed a rate of 500 M³cf/day. During light load periods at night, on the other hand, deliveries from the line may fall as low as 100 M³cf/day. The difference between the 100 M³cf/day rate and the approximate 300 M³cf/day rate being delivered into the line at Blythe is the gas that is being packed into the line during the night period for withdrawal during the subsequent day. The total withdrawals from the line during a 24-hour period are approximately 305 M³cf/day, the same as the amount of gas which is put into the line during the corresponding period at Blythe. The storage designed into the line is equivalent to a holder capacity of the order of 52 to 75 M³cf." W. C. Mosteller, "Transmission Line Type Storage—Capacity Calculations and Costs," *Proceedings*, Natural Gas Department of the American Gas Association (1950), p. 35.

allocated to the firm customers and none to the interruptible customers. This argument completely ignores the fact that the gas sold to the interruptible customers could have been stored and used to meet peak firm requirements with a consequent decrease in the necessary production capacity. In the example used, the interruptible sales should be allocated to the production capacity cost in excess of the capacity cost for 12,000 cubic feet per hour minus any savings in storage capacity cost.

In order not to have to construct another diagram, assume that Figure 8 represents an annual load curve, with gas in cubic feet measured on the y axis and the twelve months of the year on the x axis. Assume further that the economic equilibrium value for storage capacity is 36,000 cubic feet of gas; therefore, the company has a peak rate of delivery from the pipeline of 14,000 cubic feet per hour. Assume that the storage is used to meet the peak requirements during the heating season.[46] Under these conditions, any interruptible sales made during the heating season when gas is drawn from storage will add either to production capacity cost or storage capacity cost and such increase in capacity cost is correctly allocated to interruptible users.

On the present assumptions (i.e. storage capacity of 36,000 cubic feet), production capacity cost would be allocated to the time period marked as 9 A.M. to 9 P.M. in Figure 8 and the storage capacity cost is allocated to the time period indicated as 9 A.M.-1 P.M. and 5 P.M.-9 P.M. Any interruptible sales made during the period indicated as 9 A.M. to 9 P.M. in Figure 8 would add to capacity cost and therefore should be allocated to the interruptible customers. Assume that the company represented in Figure 8, dispenses with storage capacity, increases production capacity from 14,000 cubic feet per hour to 19,000 cubic feet per hour and sells the "excess" gas over the "firm" customer requirements to customers under interruptible service contracts. Under these conditions, the interruptible customers should be allocated the production capacity cost for 5,000 cubic feet per hour minus the storage capacity cost for 36,000 cubic feet of gas.

The Federal Power Commission's Treatment of Interruptible Service. The Federal Power Commission has argued very strongly against assigning all of the costs that vary with capacity to capacity cost; the Commission argues that part of these costs should be assigned to energy cost. Their argument appears to be based on a refusal to accept the results of a supposedly economic allocation based on peak responsibility, an allocation which, in fact, is incorrect because the consequences of gas storage have been ignored. The Commission is not alone in ignoring the implications of gas storage for the allocation of capacity cost.[47]

In the Seaboard Case the Federal Power Commission stated its views on the problem of cost allocation at some length. The following quotation presents their position:

[46] "The use of . . . . storage projects is associated essentially with the delivery of gas during the heating season that normally extends over a period of from three to four months." Roy A. Wehe, *Cost to Serve Study*, Southern Counties Gas Company, Exhibit 11 in Application No. 33341, before the California Public Utilities Commission, 1952, pp. 12-13.

[47] The only method of allocating capacity cost specifically designed for gas transmission and gas distribution utility companies, the Capacity Credit Method, underestimates the factor of storage. "Since gas and electricity cannot be stored in appreciable quantities . . . a public utility has to produce, transmit, distribute, sell, measure and deliver its commodity or service to all its customers simultaneously and, at any time, in the quantities required by them." Hans E. Nissel, "New Capacity Credit Method Provides for Allocation of Costs in Gas Systems," *American Gas Journal*, October 15, 1953, p. 16. Nissel acknowledges that underground storage somewhat alleviates the situation for gas transmission systems. In his method of allocating capacity cost, Nissel appears to lump storage capacity cost with other capacity cost—a procedure which, in general, does not give an *economic* allocation of storage capacity cost.

"A pipeline would not normally be built to supply peak service, that is to say, service on the peak days only. We know from our administration of Section 7 of the Natural Gas Act, which involves the issuance of certificates of public convenience and necessity, that pipe lines are built to supply service not only on the few peak days but on all days throughout the year. In proving the economic feasibility of the project in certificate proceedings, reliance is placed upon the annual as well as the peak deliveries. Stated another way, the capital outlay for the pipeline facility is made—and justified—not only for service on the peak days but for service throughout the year. Both capacity and annual use are important considerations in the conception of the project and in the issuance of certificates of public convenience and necessity. Both capacity and volume, therefore, are what are known as cost factors or incidences in respect to the capital outlay for a pipe-line project. It follows that reasonably accurate results can be achieved only by allocating the fixed expense flowing from the capital outlay to both operating functions, viz., capacity and volume.

"The soundness of the foregoing conclusion will be made clear by a simple illustration. If fixed expenses are assigned wholly to the demand or capacity function, then gas service which is interrupted on peak days will not share in any of the fixed costs. Conceivably under such an allocation large quantities of natural gas could be sold to industrials 360 days of the year and interrupted on five or so peak days, and such gas would bear none of the costs incurred in constructing the pipe-line facility. In other words, under a strict application of this theory, the interruptible service would not bear any depreciation expense, return, income taxes, or any part of the other fixed expenses associated with the capital outlay. This would be so even though the capital outlay made the interruptible service possible by providing the means of transportation used extensively by such assumed interruptible service."[48]

The Commission's argument is not in accord with an economic allocation of capacity cost. It is clear from the illustration given by the Commission in the above quotation that they do not consider the implications of gas storage on capacity cost allocation. The Commission's belief that interruptible customers should be allocated some capacity cost is economically correct under the operating conditions found in the gas utility industry, not for the reasons they give, but because gas is stored. Their reasoning, however, appears to be based more on what is "reasonable" or "equitable" than on an economic allocation of capacity cost.[49] In the Seaboard case they arbitrarily allocate 50% of the capacity costs to energy cost.[50]

At another point in the same opinion, when discussing a clause which limited the amount of gas available for use as boiler fuel, the Commission writes "that such a restrictive provision is in furtherance of the existing practices of the operating companies in the Columbia Gas System, including applicants, and is principally required to permit the storage operations which are necessary to the most satisfactory operation of all of the Columbia Gas System companies."[51] As it has been pointed out previously, production capacity cost is not correctly allocated only to the peak hours or half dozen peak days of the year, under an economic allocation of capacity cost, where storage of gas is economically feasible. If the production capacity operates at a 100% load factor, production capacity cost should be allocated to all units of output produced during the year; if the production capacity operates at less than a 100% load factor during some parts of the year, production capacity cost should be allocated to the output consumed during the

[48] Federal Power Commission, *Opinion No. 225, in the Matters of Atlantic Seaboard Corporation and Virginia Gas Transmission Corporation*, Docket No. G-1384; *Atlantic Seaboard Corporation*, Docket No. G—1175, pp. 20-1.

[49] "To reiterate—and emphasize—fixed costs or expenses are incurred for both peak use and annual use in respect to both demand and volumetric functions. They are important cost factors in respect to both services. To achieve a reasonable equitable result—they must be apportioned to both services." *Ibid.*, p. 23.

[50] *Ibid.*, p. 23.

[51] *Ibid.*, p. 43.

time when the capacity is fully utilized. Storage capacity cost is, of course, allocated to all units of output consumed during the time when gas is drawn from storage.[52]

The allocation of production capacity cost to the output consumed during the period when production capacity is fully utilized applies to interruptible gas sales as well as firm sales because during this period gas sold to interruptible customers could equally well go into storage to be used later to meet peak firm requirements.[53] Therefore, interruptible gas sales during the peak season of the year, when gas storage is economically feasible, are correctly allocated a portion of production capacity cost under an economic allocation of capacity cost as developed above for gas utility companies. The precise allocation of capacity cost to interruptible customers will, of course, vary from company to company depending upon the specific circumstances but the general principles for such an allocation are clear.

To summarize: first we developed an economic allocation of electric capacity cost based on the allocation to each unit of output the long-run marginal or incremental capacity cost of producing that unit of output. The amount of capacity an electric utility company is required to maintain to meet consumption demands depends on the annual maximum rate of consumption or "demand." Therefore, in an economic allocation the capacity cost should be allocated to the output produced and consumed during the peak hours, with relevant hours being the expected peak hours, not the peak hours of some past period. Various alternative methods of allocating capacity cost were compared with each other and with an economic allocation. Not only do all the alternative methods allow consumers to reduce their capacity responsibility by making consumption shifts that are detrimental to the system but all the alternatives seriously under-allocate cost to the output produced and consumed during peak periods. If the demands for electricity and gas have any responsiveness to price, the use of the alternative methods of allocating capacity cost where rates are based on cost brings about a decreased use of existing facilities, an increased investment in new facilities and a more inefficient use of the new plant than would exist under an economic allocation of capacity cost. The same situation exists wherever rates are based on value of service and the rate schedules do not reflect the essential variation between peak and off-peak periods of use. The factor of gas storage, generally ignored, complicates the situation for gas utility companies but the subsidization of peak use by off-peak use is present, leading to less efficient use of plant than would exist if rates were based on an economic allocation of cost.

[52] In an article dealing with adequate rates for interruptible industrial gas, Paul E. Peacock, Jr., incorrectly allocates all of the cost of peak-shaving to domestic customers and none to firm commercial or industrial customers. Peak-shaving is the practice of manufacturing gas locally to meet the peak rate of consumption and thus to reduce a company's "demand" charge from the pipeline. If no interruptible sales are made while peak-shaving capacity is being used, no peak-shaving capacity cost should be allocated to interruptible service. However, peak-shaving capacity cost should be allocated to all output consumed during the period when the peak-shaving plant is in operation, including commercial and industrial customers, because gas sold to these customers could be diverted to meet domestic requirements and vice versa. Paul E. Peacock, Jr., "Advantages of Adequate Rates for Interruptible Industrial Gas," *American Gas Journal*, January 1954, pp. 16-7, 33-4.

[53] The importance of gas storage is made very clear by S. A. Chadwell of the Columbia Gas Systems. He writes: "Before the development of underground storage, it was necessary for the dispatchers of most companies to follow the hourly fluctuations in market demand by adjusting pressures in the producing fields. This technique is still used in a minor way, but the dispatcher now depends on the manipulation of underground storage to absorb the major part of the market fluctuations." He adds: "The use of underground storage, particularly in the Appalachian area, has made it possible to construct and operate many pipelines from the Southwest which otherwise would not have been economically feasible. Its value to these long line companies results from having placed them in a position of purchasing gas and operating their facilities at near 100 per cent load factor." S. A. Chadwell, "Pipeline and Network Gas Dispatching," *American Gas Journal*, April 1954, p. 27.

Frank Lloyd Wright: His Contribution to Our American Culture

By RALPH A. WALKER*

Introduction

Frank Lloyd Wright, architect extraordinary, is a product of Wisconsin. In a peculiarly significant fashion he epitomizes the spirit of the founder of this journal, Richard T. Ely, as well as another Wisconsin contemporary, John R. Commons. These three scholars were endowed with an exceptional ability to venture into unknown and untried experiments—in the social as well as in the physical sciences—and in so doing raised the eyebrows of both colleagues and critics. Perhaps it is of such stuff as this that the creative mind differs from the pedant. Many of the once startling theories and innovations of their earlier days are now accepted as standard and even basic. Was it not ever thus?

This urge to explore new ideas is synonymous with education. The governing body of the University of Wisconsin called it "that fearless sifting and winnowing" after truth and in 1894 emblazoned their words in stone for posterity on the facade of the main building on its campus. That Frank Lloyd Wright is an educator who has pioneered and in fact fashioned the pattern of an American architecture is obvious. The professionals in allied fields—land economists, city planners, engineers—seek to analyze the impact and the environmental influences of the hundreds of Wright-designed homes, factories and churches. The sociologists, too, are having their innings with a piece on "What It Means to Raise a Family in a Wright-designed Home." But the legal acceptance of this educational quality of Mr. Wright's work has not yet been recognized in Wisconsin. The Wisconsin Supreme Court ruled that the Frank Lloyd Wright Foundation School at Taliesin, Wisconsin, does not constitute a tax-exempt educational institution in accordance with its interpretation of the current state tax laws. This decision caused a protest from Wisconsin citizens as well as architects and educators throughout the country which was formalized in a testimonial dinner held at the University of Wisconsin in February 1955. Among those who spoke in tribute to Mr. Wright and his work—in addition to the Governor of the State, Walter J. Kohler, and the Emeritus Dean of the College of Letters and Sciences, George Sellery— was Ralph Walker, well-known and successful architect of New York City. This is his tribute. (MARY E. AMEND, *Editor*)

* Architect, New York City

HIGH over a doorway leading to the University of Aberdeen in Scotland there is this motto: "Thai Half Said. Quhat Say Thai? Let Thame Say." If I might take any theme for my few remarks it is in praise of this old 16th Century idea—one given so much authority by that proud and independent university.

It is not my purpose here to discuss the laws of the State of Wisconsin or the acts of jurisprudence which govern the decisions concerning them; nor shall I speak of the justice, or otherwise, of the taxes within the state—all of these surely are more intimately the concern of its citizens—except perhaps to say that all laws made by man are open to a dissent in opinion and are liable and free, in sure ways of justice, to make a unique exception by liberal interpretation; and that, as far as taxes are concerned, all of us have a growing and gloomy specter at our backs, one who reaches over our shoulders to clip off great morsels of our efforts—so much so that finally many of us, perhaps unfortunately considered successful, spend more time working for our several and multiple governments than we do for ourselves, for our families; and, doubly unfortunate for our employees—whose security no longer belongs to them, to us, or in final analysis, to almost no one. *This sad story accompanies the age of science and its plenitudes.* Wisconsin, as well as the rest of the world, is continually faced with the necessity of making a positive choice between the forces of desperation and man's long search for the beautiful and the good.

Sophocles said: "Many are the wonders of the world, but none more wonderful than man who conceives and carries out such plans." It is in tribute to a man, an architect, a creator, that I would speak; in tribute to Wright's long and successful search for beauty. *One attains stature, even in our neurotic world, by being one's own self: in developing one's own soul into a comprehension of its potential greatness.*

It is obvious that we live in a world of very few heroes indeed. (A hero, in my definition, is an individual who has resisted the stereotype of his time.) In a world which more and more demands anonymity, Wright stands out as a true individual, more and more as a most vehement protest against the rise of a completely standardized life. I need not say that his concept of Usonia, his many ideas developing the "Broad Acre" city, are the very opposite of the serf state. Independent,—there are those who think of him as independent as hell—as a pioneer he approaches the magnitudes of a hero; but it is not just as a pioneer that Wright stands: no, but as a great mountain, unyielding against the grey skies of homogenized humanity. I have said elsewhere that he asked a drab society to compromise with him to make a better society; one not bound by fear; one not bound inevitably to an engineer concept of the machine. Wright, while understanding the forces of structure, has always made the architectural concept the master, so that structure alone has had but small meaning—other than that which man throughout all time has meant it to be, and that *is*, the bones which are to be covered with the flesh that makes for a living beauty; not ever for the monotony of skeletonized death. Therefore as you look at Wright's works you find complete harmony of purpose, of structure and of fair covering; these raising each work to a position of unity, unity which, regardless for whom designed, is a marked characteristic of all his architecture. Nor has this been a late phenomenon because it is obvious that the young Wright working far into the night foresaw, not as a conscious sooth-

sayer perhaps, but as a creator, the architectural expression for a new world—a world which about him and for some years had been in accelerating revolution. But in his elder days, having kept the faith to that vision, he reaps a self-satisfaction which rightfully belongs to a man who, throughout his life, has maintained the integrity of his own soul.

It is the sure thing to be noted that throughout bad times, through small and high success, through great tragedy or small happiness, through early denial, and finally through the growing and immense respect in which we now hold him, he has always maintained the integrity of his soul's purpose. Call it courage, even obstinacy if you will; it is that quality which, to all of us, will make him stand out for our emulation, and make him great for all time.

If one had been omniscient in 1890, one might have looked with wonder at the great divisions in architectural design which were opening up at that time; the rise to new heights of an eclecticism of the classic past—who would have thought of it as the throes of imminent death; the beginning, a slow birth of a new way of creating architecture, one not barren but rich in its knowledge of form and of materials, but which is still unfortunately, all too slowly being accepted; and again, that great burden now placed upon an undeserving world and which, evidently all about us, is proving a dead-end of architecture, i.e., the development and acceptance of the engineered-inspired building and city; a dead-end of architectural and city design which finally *weeds up* in such projects as the Fort Dearborn plan for Chicago—surely a fatal flowering of highly glazed mediocrity.

Does the engineer know best? Are there no indications that monotony is on the wane or that a richness in imagination is not a thing of the future? Is our society driving gradually, but irresistibly, toward an ant civilization? Are the Mies bones of concrete and steel the end of the restless and amazing urge for continuing creation that has raised man above the brute animal from which he sprang? De Tocqueville, in his famous book on democracy dreaded the then evident failure of a democratic society to develop the highest moral and intellectual qualities in man. He believed the worst threat to mankind within the democratic system was that mediocrity would not only be encouraged but be enforced.

We must therefore hold as precious a man like Wright who denies the mediocre, who acclaims the individual search for beauty, who strives for meaning within our democratic way of life; for, as has been said, of all forms of social existence, democracy, with all its faults, is still the best.

The question we must answer is whether our civilization, with all the plenitude of our resources as developed under our democratic form of government, can do other than supply the bare necessities of life so that there will be little left over for the true amenities which have so often enriched the past. If there are no amenities then certainly we are to be faced with an architecture quite different from that heretofore known; and the factory, a graceless thing in appearance, so prim and spare, so jail-like, will carry on as it has for the past seventy-five years, increasing in endemic manner to blot out the character of every other type of architectural purpose. In this immediate past there have been few bright spots architecturally, except Wright. Eclecticism of one sort or another has been all too general and fashion has succeeded fashion to such an extent that the city especially has lost all sense

of harmony, all sense of those qualities of "locale" which have given such welcome savor to past civilizations.

And why has this happened? Perhaps because the architect has admitted the accusation that architecture is something which is added on rather than what it is, i.e., the truly fundamental and basic reason for shelter, what Wright calls the organic quality—that quality which, were we to heed its existence, would give our architecture the importance our intelligence, if not our sentiment, even now indicates is possible.

In times past the structure has always been the servant of the creator: human need, whether it was merely the shelter requirements of poverty or the ecstasy of high religious fervor, has used architecture, meager or grand, to pleasant purpose. Never before was there expressed the fear of life, the fear of history, the fear of enrichment of surface; and as a result there has finally developed a habit of thought which stupidly refuses to encourage individuality. This is a time when all societies not only encourage depersonalization, but enforce it, and so we find a steady growth of uniformity; and the little imagination there is, is used by designers to become pioneers in imitation. Man tends to become a scientific insect.

Probably the most remarkable thing about Frank Lloyd Wright has been the amazing fertility of his mind. It seems throughout his life that each problem presented to him has called forth new resources, new ideas which flowered as fresh blooms of genius. Nor has age greyed the green thumb in his architectural garden of life. Quite the contrary, each design comes to life with shining character, one gently held in the bud of youthfulness; and yet all of them breathe of this age, our age as it might be, no monotony, no mediocrity, no submission to the meagerness of dead-ends, each design hopefully saying, here in this charnal house of modernity, is a symbol of what the will to create may bring forth for the happiness of man.

It is this symbol of life, of the architect as the creator, of a man who not only believes he can mold his civilization, but who actually tries to do it, of a man who refuses to abdicate to the ignorant mores of the "group spirit," and "of collective thinking," and of "integration," all modern terms denoting submission in a fear for security—that we admire in Wright. If heads are to be counted, if heads are to be cracked, Wright stands ready to take whatever his creations may bring him. His is the pride of a man whose integrity forbids him to bow in fear of being eliminated.

Wright belongs to a greater group, one famous for invention, a group which throughout the ages, individual by individual, has added to the ever present ape within the common herd the attributes of man the thinker, and finally to the best thought we find now in our own civilization. Little by little the creative man (all too few in number) has added to the enrichment of our ways of thinking. of living, of achieving greatness.

Fearless in facing life, creator for its betterment, persisting with a high integrity of soul, Wright lives—and will long endure in memory. Wisconsin evidently takes to its heart, in generous acknowledgement, the greatness to which it has given birth. Let it never be said that Wright was without honor among his own people. May I say, finally, that some future day, without doubt, Taliesin will be made a shrine by the people of the State of Wisconsin and of our nation.

Milestones and Memoranda on the Work of Frank Lloyd Wright*

(*Editor's Note.* In the late summer of this year of 1956 Frank Lloyd Wright made some public utterances on the skyscraper. He seemed to be saying that he disapproved of its present form and function, especially when it crowds in on the central core of the city and creates ugly skylines. He extended his comments to describe what he would do to make the skyscraper both durable, functional and, equally important to man, esthetically pleasing. He would make it sky-high. Again he caught the attention of the cautious, the curious—and the critics. He has put in writing a Specification for the Mile-High Illinois and it is included below in these thumb-nail descriptions of some of the significant works of Frank Lloyd Wright through over a half a century of progress in the development of an American architecture. From the original Prairie House to the Mile-High Illinois the fact is clear that the once-startling ideas and concepts have taken root in the urban and rural planning redevelopment scene. Who will gainsay the architect's prophecy on the Mile-High Illinois? It is put into the record here for 2006 A.D. scholars to evaluate along with the *institutional economics* of John R. Commons, and that of the founder of this journal, Richard T. Ely, when he expounded on "*under all, the land.*"

1911

SPRING GREEN, WISCONSIN *Taliesin or "Shining Brow"*

Home and studio, built on the brow of a hill rather than the top. It is a house of the North, with sheltering roofs for winter snows and wide overhangs to catch summer breezes. Masses of native limestone were laid up for the first time in courses just as they were discovered in the quarries. Plaster inside is dyed a tawny gold, and outside left its natural sand color. Ceilings of many interiors follow the roof planes above, enriching the space within. Outside, the cedar-shingled roofs are left natural to turn a silvery-gray. Various balconies offer commanding views of the valley and lakes below and the Wisconsin River beyond. Terraces tie the house into the garden courts behind. Quiet and harmonious, Taliesin belongs to the landscape of southern Wisconsin.

1938

NEAR PHOENIX, ARIZONA *Taliesin West*

The desert terrain was the result of geological cataclysm—angular mountains, jagged rocks, unique, even prehistoric plant growth, brilliant colors—in strong contrast to the soft rolling hills and lush green of Wisconsin. In 1927 there was built "Ocatilla." There in it canvas ceilings were first used for their translucence and economy. In 1938 on Maricopa Mesa, overlooking Paradise Valley near Phoenix, there was started a great building which had neither precedent nor successor: Taliesin West. Colorful stones found on site were placed in wooden forms and held together by concrete. Sand for the concrete was dug out of nearby washes. When the forms were removed they revealed the so-called "rubble-walls." Superstructure and roofs were redwood and white canvas, giving the camp the appearance of a great ship a-sail on the desert. Taliesin West is so integral a part of this timeless environment that it is timeless itself. Work on the camp is still in progress, for Taliesin West is really in the nature of an experiment in materials and construction methods suitable to the desert. This participation in building construction is a vital part of each apprentice's architectural education while living at Taliesin.

1896

TALIESIN, WISCONSIN *Romeo & Juliet*

This windmill, towering above surrounding trees, is the combination of a working principle and artistic expression. Vertical metal straps are anchored in a deep stone foundation as the roots of a tree in the ground. The wooden superstructure, bolted to these rods, makes the whole structure as impregnable as a barrel. Romeo, the prow, faces the direction of strong winds; while the observation tower, Juliet, clings safely alongside. This early engineering-architecture has long outlived the doubting valley residents who, after each storm, would come to their doors to see if the tower were still standing. It is—after sixty years—upright, slender and graceful as the day it was built.

Project: **1897**

CHICAGO *Luxfer Prism Building*

A revolutionary facade of sixty years ago—contemporary today.

* First presented as captions for the Work of Frank Lloyd Wright exhibited in Chicago, October 16, 17, 18, 1956.

1904

BUFFALO *Larkin Building*

The first great Affirmative Protestant against the superfluous and meaningless ornamentation of the period and abuse of materials. The first Great Pioneer—in simplicity of line, integration of form and purpose and such innovations as the first complete air conditioning, metal furniture, wall-hung water closets, complete plate glass doors and windows, etc., etc., Understanding the industrial needs of our century, this office building rendered the time and place of modern man as a work of art. It emerged half a century ahead of its time, and is now destroyed.

1906

CHICAGO *Unity Temple*

The first concrete monolith to come from the forms as architecture complete. The work was cast in wooden forms or boxes—and the building bears the impress of that technique. This plan first began the destruction of the box, and the emphasis of interior space as the reality of the building. This has been carried on.

1908

RIVERSIDE, ILLINOIS *Avery Coonley House*

This became one of the first great "Prairie Houses" planned with the concept of zoned living: all functions separated, individualized, yet brought together by galleries and bridges. Living areas and bedrooms are on the second floor, bringing the basement out of the dark and level with gardens and pools. The unit of design is the human figure, a sense of scale totally new to the language of architecture. Horizontality, low roofs, garden walls and exterior planting boxes keep the home in sympathetic harmony with the midwest prairie. Colored tiles set into cement plaster, ornamental windows and copper pergolas geometrically complement and enrich the plan. All furniture, rugs and lighting fixtures were planned to make each part relate in proportion and design to the whole.

1909

CHICAGO *Frederick Robie House*

This early work abolished the idea of architecture as a series of boxes punched with holes for doors and windows. In its place came these new concepts: Interior space as the significance of a building, with exterior form as a natural result. Purpose and design treated as one. Each material considered for its own properties—stone as stone, brick as brick, wood as wood, etc. The Robie House was one of the first examples of a new principle of domestic architecture. The game room and play room, previously a damp hole in the ground, became the ground floor. The living areas were raised to the second floor. Low pitched roofs with wide overhangs sheltered large areas of glass and eliminated the awkward attic, presenting a new horizontal "streamline" destined to revolutionize the entire grammar of twentieth century architecture.

1914

CHICAGO *Midway Gardens*

A complete ensemble of cabaret and dining gardens, the Midway Gardens was the first instance since the thirteenth century when architecture, sculpture, painting, music, ornamental glass and landscaping came together under the design and supervision of a master architect. The outdoor orchestra shell had its origins here. But the success of this revolutionary handling of materials and crafts was ignored by the public. Like the Larkin Building (Buffalo, 1904) Midway Gardens remains for us today only in drawings and photographs: a brilliant exposition of American culture built at a time when culture was sought across the Atlantic Ocean.

1916

TOKYO *Imperial Hotel*

A richly ornamented building—in harmony with the culture of Japan though embodying structural principles new even to the West. (This hotel was commissioned by the Imperial household after its delegation had conducted a world-wide search for an architect.) The intrinsic stability of the cantilever—as the waiter with his balanced tray—and the continuity of steel in tension—as the interlocking of the fingers of the hands—were the chief structural innovations employed by Mr. Wright to make the wide-spreading building earthquake proof. The disastrous temblor of 1923 proved the genius of this scheme and the hotel still remains the most outstanding building designed by a foreign architect in Japan.

Project: **1916-1920**

WATER TOWER SQUARE, CHICAGO *The Glass and Metal Skyscraper*

A triumph of light and air over the dark granite jungles that comprise our cities.

This project of forty years ago for the National Life Insurance Company heralded a new understanding of the Twentieth century miracles that might be worked by reinforced concrete, glass, sheet metal and industrial methods. The tall building need no longer be a steel-framed box: one box upon another. The cantilever principle—steel in tension— successfully proven in the Imperial Hotel (1916) in this building kept the support of the floor slabs in central pylons and left the exterior free of heavy masses, a pendant screen of sheet copper and glass. A forerunner of the mile-high "Illinois."

Project: **1922-1929**

NEW YORK CITY *St. Mark's Tower*

Intended to rise in a group of three out of a small park, this apartment tower is another expression of central-core cantilevered floors and pendant screen walls. All rooms are exposed to sun and air. The plan, a quad, is arranged to provide for the complete privacy of each unit-duplex. Centralization, essential in the large city, is here presented by imagination rather than standardization. Another precursor of the mile-high "Illinois."

1936

MADISON, WISCONSIN *The First "Usonian" House*

The suburban family of moderate income in America had long been forced to accept the commonplace box for their home. This house, the first of a type to be known as "Usonian," offered a new solution both economic ($5,500.) and artistic. A concrete mat floor, containing heating coils (the first instance of floor heating in America) eliminated expensive excavation for a basement. The overhanging roof sheltered the house and became the carport, eliminating the garage. Walls and partitions were fabricated with a plywood core faced on both sides with insulation paper and boards and battens, eliminating wasteful stud-and-plaster construction. The plan in general is an "L" opening onto an enclosed lawn and garden. A central core contains heating plant, utilities, kitchen and bathroom, while one wing serves as the living area and another for bedrooms.

1936

RACINE, WISCONSIN *Administration Building—Johnson Wax Company*

Elimination of the cornice. Slender columns, fanning out at the top to form giant "lily pads," become the ceiling; the spaces in between, the skylights. Brick walls therefore become non-supporting screens. The corner where the wall meets the ceiling is here—for the first time—replaced by translucent glass tubing. Emphasis is on light from above in this building situated in an industrial environment.

1936

BEAR RUN, PENNSYLVANIA *Falling Water—House for E. J. Kaufmann*

Twenty years ago a dramatic site in the Pennsylvania wilderness became the inspiration for the first dwelling in reinforced concrete utilizing the cantilever principle. Projecting balconies, counterbalanced by heavy masses of native stone, carry domestic living out over the falls and into the surrounding trees.

1937

RACINE, WISCONSIN *Wingspread—House for H. F. Johnson*

Last of the great "Prairie Houses." Expensive living in America need no longer be interpreted by a setting of grandomania. This house for the president of the Johnson Wax Company was designed in such a way as to bring domestic living closer to the feeling and understanding of a beautiful natural setting. Expansive in its scale of living, nonetheless, the attention of this house is focused on a great midwest prairie—with all the joys that accompany living close to the change of seasons in southern Wisconsin.

Project: **1939**
Under construction: **1956**

NEW YORK CITY *Museum for the Solomon R. Guggenheim Foundation*

Now under construction on 5th Avenue, this museum will solve the age-old problem of all museums: circulation. Elevators will take visitors to the top where a spiral ramp leads them down along a gallery to the ground floor. One continuous spiral of reinforced concrete, all as one floor. Open court in the center, lighted by a glass dome above. Continuous spiral skylight affording indirect daylight on the paintings of the gallery wall. Sealed and air-conditioned. A separate unit contains the Foundation offices. When complete, New York City will have its first instance of the third dimension in architecture.

1940

LIBERTYVILLE, ILLINOIS *Lloyd Lewis House*

On a site beside the Des Plaines River, this house was built up off the damp ground to afford an intimate view over the water's edge and amongst the dense forest growth. Natural cypress is left to weather a silvery-gray, and common brick is used for masonry areas. All furniture and decorations were executed as planned by the architect, making a most harmonious and complete interior.

Project: 1940

HILO, HAWAII *Martin Pence House*

Intended for a hot-weather climate, this was the first house designed with no interior corridors: each room is a separate unit that opens onto a large loggia exposed to a pool on one side and a walled garden on the other.

1940

BLOOMFIELD HILLS, MICHIGAN *Gregor Affleck House*

The living area is carried out over the edge of a ravine by a great cantilevered balcony, thus taking complete advantage of the steep slope rather than destroying it by costly excavation.

1940

LAKELAND, FLORIDA *Anne Pfeiffer Chapel—Florida Southern College*

Screen walls of perforated concrete block, filled with colored glass, bring colorful and subdued patterns into the darker recesses of the room, while a central lantern tower pours light in from above. This chapel is one of 8 buildings completed on the Wright-designed campus.

Project: 1940

WASHINGTON, D. C. *Crystal Heights*

The cantilever principle of construction—from within outward. Five parking terraces surmounted by a spacious park serve as the plaza from which rises a group of 15 interlocking hotel towers. Lobby, offices, shops, and restaurants face gardens on the plaza level and a cinema at the main apex of the site is accessible from both the street and within the hotel. All combine to render a perfect solution for travelers to our congested cities.

Project: 1946

DALLAS, TEXAS *Rogers Lacy Hotel*

Translucent walls of fiberglass (a material having the insulating value of an 8-inch brick wall) are hung from cantilevered floors and rise in the form of a crystal tower above a great mass containing the hotel's public rooms. A concrete pylon, extending from the foundation up through the roof garden, stabilizes the tower and carries ample air conditioning and all utilities. This building was planned around an interior open court, eliminating interior corridors. Gardens, pergolas and fountains are incorporated in the sun-lit court as a relief from urban monotony.

Project: 1946

SAN FRANCISCO, CALIFORNIA *V. C. Morris House*

Seaside home, clinging to a cliff. Reinforced Concrete. Earthquake-proof.

1947

RACINE, WISCONSIN *Research Laboratory Tower—Johnson Wax Company*

The first proven tap-root foundation. (Although this idea appeared in the St. Mark's project, 1929.) Cantilevered floor slabs, screen walls of glass tubing, brick parapets supported by the floor slabs. Sealed and air-conditioned. The tap root foundation principle was also executed in the Price Tower, and is planned for the mile-high "Illinois."

Project: 1947

HOLLYWOOD, CALIFORNIA *Sports Club for Huntington Hartford*

With imaginative use of the cantilever, the elements of a 20th century good-time place were dramatized for a wooded hill-crown close to downtown Hollywood. Three discs spring from a staunch central mass to uphold the glass-domed dining room, salon and cabaret and overlook swimming pool and championship tennis court ranged on bowl-shaped projections which stem from sweeping terraces at the base. The club was planned as the recreational unit of a complete hotel to have been built in the canyon below.

Project: 1948

PITTSBURGH, PENNSYLVANIA *Pittsburgh Point Bridges*

A great concrete mass counter-balances the weight of road beds via an artistic application of the suspension principle: tensile steel cables, woven into a rich pattern. Pedestrian walks and gardens elaborate this scheme.

Project: 1949

SAN FRANCISCO, CALIFORNIA TO OAKLAND *Bay Crossing*

Gun-concrete, sprayed onto pre-stressed steel eliminates expensive formwork and the continuous maintenance of an exposed steel bridge. The roadbed is cantilevered from central supports, making bridge and pier one unit. In the center the bridge rises for a 1000-foot leap over the main ship-channel. At the top of the arch the roadbed divides to pass on either side of a suspended park.

1949

MADISON, WISCONSIN *Unitarian Church*

The steeple abandoned as a cliche, a sweeping roof over a triangular plan combine to give the feeling of shelter, worship and aspiration.

1950

SAN FRANCISCO, CALIFORNIA *V. C. Morris Shop*

An arched entrance set in a wall of tawny brick invites the passer-by into this intimate gift shop to explore merchandise exhibited charmingly in a setting of ramps and small alcoves.

1952

SCOTTSDALE, ARIZONA *David Wright House*

Situated in an orange grove, this house presents an ideal way to live in the southwest. The main living area and bedrooms, rather than resting on the hot desert floor, are raised on broad columns and approached by a wide ramp (in itself a garden) to afford a view of the surrounding mountains.

Project: 1952

ACAPULCO, MEXICO *Raul Bailleres House*

A boulder-covered site overlooking the Bay of Acapulco gave rise to the general form of gunite arches and domes. The tropical climate permitted living areas to open out as pergolas and loggias. Ramps connect the house to a pool and a beach below. Unlike the customary "Period" or glass boxes designed for luxurious resort living, this home—despite its great size—blends with and enhances the site.

Project: 1953

SILVER SPRING, MARYLAND *Llewelyn Wright House*

Circles and ovals produce a dramatic use of the cantilever to bring domestic living out and into the surrounding trees on this steep hillside, much in the same fashion as "Fallingwater" of 1936.

Project: 1953

GRAND CANAL, VENICE, ITALY *Masieri Memorial*

Here in this city where Asia Minor and Europe met and traded for centuries is the projected design for a library and dormitory to accommodate 12 students. Reinforced concrete pylons driven down into the ocean bed stand irrespective of the movement of tides and sands. These pylons rise to the surface and into the sky to become the vertical supports that carry reinforced slabs. All walls and glass are screens. Pylons and parapets are faced with blue-veined white marble. Arabesques of geometrical design stamped into copper grace the spire of light running up the side of the building and above the belvedere on top. The entire feeling of this charming little jewel is one of beautifying a sound twentieth century structural method in such a way as to blend harmoniously with Byzantine Venice.

1954

PHILADELPHIA, PENNSYLVANIA *Beth Sholom Synagogue*

A new and inspiring form is here given to the Hebraic faith, with the building and all its details related to Hebraic symbolism. The synagogue itself—Mount Sinai. A tripod of steel beams, faced with stamped metal, rises to an apex one hundred feet above the floor. This tripod carries a double screen of translucent plastic panels, providing insulation and light for the great hall inside.

1955

BARTLESVILLE, OKLAHOMA *H. C. Price Tower*

Urban centralization is here shown as a beautiful feature of the vertical line: tall, slender, light. Sun and air for all the occupants. Flowers and plants on projecting balconies and roof make the building a garden city—a tree rising in a park. Completed last year, this tower is a further development of the St. Mark's scheme of thirty years ago. Eight duplex apartments occupy one quadrant, while the rest is office space. Vertical and horizontal louvers protect the glass screen walls against the hot Oklahoma sun. Another forerunner of the mile-high "Illinois."

1955

SCOTTSDALE, ARIZONA *H. C. Price House*

Out of the Arizona desert grows this concrete block house, framing a beautiful view of all the surrounding mountains. The point where the block piers meet the roof is exposed by a slender metal column, giving the effect of a roof that glides over the concrete mass—yet never touches it. Pools and gardens are arranged to offer a cool, pleasant relief from the hot, arid desert.

Project: 1955

DALLAS, TEXAS *The New Theatre*

Seating, planned on the reflex, presents a three-quarter view of the stage as a relief from the customary "two-dimensional" view. A revolving circular stage allows performance on one half while scenery is arranged on the other half. A closer and more intimate relationship between actor and audience is achieved, along with all the technical facilities required of the modern stage.

Commissioned 1955

MADISON, WISCONSIN *The Monona Terrace—Civic Center*

The City goes to the lake. A civic center for the capital of Wisconsin to contain auditorium, little theatre, art gallery, sports arena, community meeting halls, offices, boat houses, parking facilities and garden terraces gives imaginative advantage to a city built on the isthmus between two lakes. Two towers carry offices for city and county governments. The entire project is designed to be built in various stages, the first of which is under active preparation.

Project: 1955

SAN FRANCISCO *Lenkurt Electric Company*

A vast industrial plant with administrative headquarters to be built on the San Francisco peninsula. Screen walls of concrete block. Dendriform colums (extension of the Johnson Wax column) for structural support. Pinnacles of copper and plastic for skylights. The entire building is raised above ground level to provide parking shelter. An artificial lagoon and park to separate the building from the Bayshore Highway. An open-air court shaded by one great pinnacle of copper louvers in the administrative section. Every effort has been made to provide the factory employee with an environment both utilitarian and beautiful.

Project: 1956

BELMONT PARK, NEW YORK *Sports Pavillion*

The spider spinning. A translucent plastic roof is suspended by tensile steel cables. Concrete bastions at either end anchor these lace-like cables. Parking, club-rooms, mutuels, restaurants and offices are located beneath the seating area.

Project: 1956

MILWAUKEE, WISCONSIN *Annunciation Church—Greek Orthodox*

From the inspiration of Byzantine form came this design for a Greek Orthodox Church. Windows of stained-glass, a dome—gilded gold inside and blue tile outside, and sculpture bring the entire building and its related arts under the direction of the architect. In plan—a cross within a cross. Simplicity, yet richness of detail proves that ecclesiastic architecture in our time can have imagination and expression suited to our materials and methods.

Project: 1956

CHICAGO *Golden Beacon: The Mile-High Illinois*

St. Mark's Tower, Crystal Heights, The Price Tower—and now the Golden Beacon. Throughout all of them the same principles at work: Central pylon carrying cantilevered floor slabs. Screen walls of metal and glass. Tap root foundation.

SPECIFICATION

1. Most stable of all forms of structure is the tripod. Pressures upon any side are immediately resisted by the other two. For general stability at great height this form is planned for the ILLINOIS to employ the new principles of cantilever—steel in suspension—as in the Johnson Heliolaboratory at Racine, Wisconsin, and the Price Tower at Bartlesville, Oklahoma. The exterior is entirely metal-faced, carried by steel wires suspended from a rigid steel core buried in light-weight concrete. The building is thus designed from inside out instead of the usual construction from outside inward. The entire structure is more airplane in character than the usual heavy building-construction. For instance, the support of the outer walls and area of the floors is pendant, and the science of continuity is everywhere else employed from inside outward. All floor-slabs are extended across the central core. All loads are thus balanced over central supports except the outer perimeter of slabs

and walls which are suspended. The science of continuity is thus employed: the type of construction similar to that of the airplane and ocean liner. This type of structure is a "natural" for either great spans or great heights. Typical weights are little more than half customary building practice.

2. This interior system of building-construction is new: Tension here involved in the ILLINOIS (upright) was first used by myself (horizontally) in the Imperial Hotel, Tokyo, Japan, 1915, and proved earthquake-proof, 1922. The same general system has now been repeated vertically—and successfully—in the nineteen story, forty-foot square, Price Tower—steel-in-tension as used years before in the Heliolaboratory. In the ILLINOIS the same principle centralizes loads over a giant core of properly designed steel fabrication cast in appropriate masses of light-weight concrete: all floor loads thus balance against each other over this staunch core but the outer portion of the slabs and the wall-screens are suspended from this core. Thus the framework of the ILLINOIS is like a tree, the horizontal floor slabs integral with the vertical core making the total structure light and rigid.

3. 528 light floor slabs are tapered hollow from the core to carry air-conditioning, lighting and appurtenance systems. These cantilevers are formed by special high-tension steel, diamond mesh reinforcement cast into light concrete slabs. Excepting vertical elevator enclosures, all exterior surface-features of this structure and certain areas of the floors are suspended from the sloping corners of the core. All outer glass surfaces are set four feet back under the metal parapets to avoid glare and afford a human sense of protection at such enormous heights as characterize the ILLINOIS SKY-CITY.

4. Transit: By atomic power, serving in series the five divisions of 100 floor heights. A group of 56 tandem-cab elevators five units high begin to load at the fifth floor, where the escalators leave off. These elevators are to be entirely independent of ordinary suspension systems. As they rise and emerge on ratchet-guides independent of the tripod into the outside air, they appear as graceful vertical features of the ILLINOIS. The entire elevator system thus rises perfectly upright to five different story-heights. Special through-service is provided to the upper stories and to the very top floor at various speeds all the way to one mile above the ground floor level. All elevators are motivated by atomic power, engines on cabs engaging ratchet-tracks, moving at various speeds much as an automobile runs on the level. Approximate speed: say, a mile per minute; appropriate automatic stop-and-go controls now familiar without attendance are provided. Additional private lifts connect various departments independently of the main elevators. Cars are set aside for non-stop emergency service. Escalators from the lower parking levels serve the first five stories. The main floor of the ILLINOIS is thus practically the fifth floor. This combination escalator-elevator service should empty the entire building within the hour by day and the various occupations by night in half the time.

5. The ILLINOIS employs again the proved system of "tap-root" foundation sloping to hard-pan or bedrock similar to the foundation of the Heliolaboratory and the Price Tower; and similar in principle to the foundation system that saved the structure of the Imperial Hotel in the 1922 temblor. To make rigidity possible at the extreme heights of the ILLINOIS, this type of foundation continues the main core into the ground to reach rock formation beneath. The foundation has available spaces within it for utilities and, owing to its tapered form and final drilling for insert of the spine, is not difficult to construct.

6. Finally, throughout this light-weight tensilized structure, because of the integral character of all members, loads are in equilibrium at all points, doing away with oscillation. There would be no sway at the peak of the ILLINOIS. A rapier with a handle the breadth of the hand set firmly into the ground, blade upright: the simile indicates the general idea of the ILLINOIS, five times the highest structure existing in the world.

7. Exterior features: Elevators, parapets, and all exposed vertical members are of gold-colored metal. Set-back of window-walls under the steel parapets to avoid glare give the building emphasis as an all metal structure.

8. Covered parking for about 15,000 cars is reached by ramps connected to one level below grade and four levels above. These

lower levels and the sub-floor parking beneath the building itself have direct access to and from escalators; there are two decks for 50 helicopters each.

In general: The ILLINOIS, divided into four parts, is entered at four points and is reached by four four-lane approaches. Fountain features and green-planted parterres are related to the building each independent of the other.

All this well done, the building will be centuries more permanent than the Pyramids.

FRANK LLOYD WRIGHT

Taliesin, Spring Green, Wisconsin

Reports and Comments

Property Taxes and Solutions to Fringe Problems: Attitudes of Residents of the Flint Metropolitan Area†

THE vexing problem of how to efficiently provide and administer urban services in metropolitan areas is too well known to require restatement. The source of the difficulty is generally traced to the fringe area, or more particularly to the multiplicity of semi-autonomous governmental units which comprise the fringe. Various explanations have been suggested for the reluctance of residents of the fringe to joint in a more inclusive governmental organization which could deal comprehensively with service requirements These include distrust of other governmental units, especially the central city; local community pride and fear of the consequences of loss of identity; vested interests on the part of local government officials; unwillingness to share the debt burdens accumulated by other units of government; and opposition to the higher taxes that would result. Needless to say, it would be of immense value to know not only the extent to which such explanations are correct, but also the conditions under which they hold and their relative weights.

This paper reports an exploration of one of the factors cited as resistances to effective reorganization of metropolitan administration, namely, willingness to pay higher taxes. But willingness to pay higher taxes doubtlessly is not an unqualified attitude. Among the many possible contingencies one of considerable importance would seem to be the kind of governmental arrangement under which higher taxes would be levied. Thus a person who favors tax increase on the assumption of annexation to the central city might not take a similar position were he to continue dependence on a township form of government. Accordingly, in the following analysis attention is focused on the relationship of willingness to pay higher taxes with preferences for governmental solutions to fringe problems. This relationship will be observed with varous demographic, residential experience, and site factors held constant.

The data for this study were obtained in connection with a larger study on the relative advantages and disadvantages of residence in different parts of the Flint metropolitan area. During the Spring and Fall of 1955 interviews were held with 670 household heads in a random sample of all residential units in the central city and outlying area or fringe.[1] The fringe portion of the sample was inflated to 3.5 times its actual proportion of the total in order to obtain enough cases for statistical treatment. Since the data on taxes apply only to household heads who own or are buying their homes, the number of cases is reduced to 542, or 81 percent of the total sample.[2]

It is necessary to stress the exploratory character of the following analysis. The data employed here were designed originally for another use, hence they lack some of the fullness desired for present purposes. For example, the information on willingness to pay higher taxes was obtained as answers to the question: "How would you feel about paying more taxes in order to have better community services in your area?" The answers were not probed for additional information on possible maximum levels or on how the increased revenues should be allocated. In connection with the data on preferred solutions to fringe problems, the question was phrased as follows:

"Many people feel that the fringe areas need such things as better sewage disposal, better police and fire protection, and so on. We have listed here a

† The authors acknowledge their indebtedness to Professors Arthur Bromage and John Hyde and to Robert Bard, T. B. Brademas, Anthony Lenzer, James Randall and Gordon Sutton who assisted in the design and execution of the survey on which this report is based.

[1] The outlying area, as defined for purposes of this study, is slightly larger than the Census Bureau concept of the urbanized area. The population contained in central city and the outlying area or fringe is approximately 75 percent of the total population of the Standard Metropolitan Area.

[2] The frequency of home ownership is 73.9 percent in the central city and 87.9 in the fringe area.

few ways in which some of these problems might be approached. Will you look over this list and tell me which one you would be most willing to support?

1. The City of Flint and the townships should work together.
2. The townships should work together to solve their problems.
3. Each township should solve its own problem.
4. The County should provide solutions to fringe problems.
5. The densely settled areas should be annexed to Flint.
6. Other (specify)."

Perhaps the most equivocal option is that having to do with the City of Flint and the townships working together. Conceivably the choices of that option might range in meaning from a casual endorsement of cooperation as a general principle to such concrete actions as contracting for water service or the extension of police and fire service. In the context in which the question was posed, however, it was clear that individuals were responding to the need for solutions to very specific service deficiencies. It is noteworthy, moreover, that more than half of the city household heads and nearly three-fourths of the fringe household heads indicated preferences for other types of governmental approaches. But, as with answers to the question on willingness to pay more taxes, the respondents were not pressed to discuss the reasons for their choices. In view of these limitations of the data the findings must be regarded as shedding a partial though nonetheless highly informative light on the question at hand.

Contrary to common belief, the fringe residents are much more receptive to the payment of higher taxes than are central city residents.[3] As may be seen in Table I, over two-thirds of the household heads in the fringe as against less than two-fifths of those residing in the central city are willing to accept increased taxes. It might be objected that "higher taxes" has different meanings for the two populations, since the reference points from which the respective judgments are made are probably at widely different levels. The data in Table II are interesting in this connection. Regardless of the amount of property tax paid in the preceding tax year, fringe residents are substantially more willing to accept higher taxes. Thus it appears that the greater willingness of fringe residents to pay higher taxes is independent of the relative amounts of taxes actually paid. A reasonable inference would seem to be that the differential in attitudes toward higher taxes reflects discrepancies in the availability or quality of urban services.

TABLE I—PERCENT DISTRIBUTION OF ATTITUDES TOWARD INCREASED PROPERTY TAX, BY PLACE OF RESIDENCE

Place of Residence	Attitude Toward Increased Property Tax		
	Favorable	Unfavorable	No Answer
Central City...	38.4	58.9	2.7
Fringe.....	68.1	28.3	3.6

TABLE II—PERCENT DISTRIBUTION OF ATTITUDES TOWARD INCREASED PROPERTY TAX, BY PLACE OF RESIDENCE AND AMOUNT OF TAX PAID IN PRECEDING YEAR

Place of Residence and Tax Paid in Preceding Year	Attitude Toward Increased Property Tax		
	Total	Favorable	Unfavorable
Central City			
Under $75.....	100.0	42.9	57.1
$75-$150.....	100.0	34.3	65.7
$150 and over.....	100.0	45.7	54.3
Fringe			
Under $75.....	100.0	73.2	26.8
$75-$150.....	100.0	71.4	28.6
$150 and over.....	100.0	62.5	37.5

When the attitudes toward paying higher taxes are sorted by preferred governmental action for solution of fringe problems the results are as shown in Table III. It will be observed that in both central city and fringe the differentials in favor of higher taxes are found among heads of households who prefer solving fringe problems through some form of joint action involving the two major sections of the metropolitan area. In the central city willingness to accept increased property taxes is most closely associated with a preference for annexation of the fringe, whereas in the fringe a favorable attitude toward higher taxes is most characteristic of respondents who prefer some form of cooperation with the central city short of annexation. On the other hand, of the central city residents who would be willing to pay higher taxes less than 18 percent prefer solutions which do not involve the central city, while among fringe residents that proportion rises to almost 50 percent. There appears to be

[3] Only 4 percent of the outmigrants from central city to the ringe give lower taxes as their reason for the move.

TABLE III—PERCENT DISTRIBUTION OF ATTITUDES TOWARD INCREASED PROPERTY TAXES BY PREFERRED SOLUTIONS TO FRINGE PROBLEMS AND PLACE OF RESIDENCE

Preferred Solutions	Attitude Toward Increased Property Tax					
	Central City			Fringe		
	Total	Favorable	Unfavorable	Total	Favorable	Unfavorable
Total	100.0	100.0	100.0	100.0	100.0	100.0
With central city	69.5	79.2	63.3	41.0	43.2	35.4
Annexation	22.7	32.7	16.1	8.2	8.4	7.6
City-twp. cooperation	46.9	46.5	47.2	32.8	34.8	27.8
Without central city	24.6	17.8	29.0	52.3	49.4	59.5
County	8.6	7.9	9.0	19.3	18.9	20.3
Inter-twp. cooperation	3.9	2.0	5.2	11.9	12.1	11.4
Township alone	12.1	7.9	14.8	21.2	18.4	27.8
Other and no answer	5.9	3.0	7.7	6.7	7.4	5.1

much greater unanimity as to the mode of solution of fringe problems among the "high tax" respondents residing in the central city than exists among similar residents of the fringe.

It is of particular interest to note that among central city household heads who are willing to accept more taxes only 18 percent, as already indicated, prefer a solution other than a joint city-fringe effort, whereas, this increased to 29 percent of those who view higher taxes unfavorably. In the fringe less than half of those who respond favorably to a tax increase prefer a solution which is independent of the central city as compared with three-fifths of those who report an unwillingness to accept an increase in taxes. Thus it is apparent that household heads in both areas who are unwilling to accept a tax increase are least likely to accept any change in the governmental approach to fringe problems, although the "status quo" finds much more support in the fringe than in the central city. This is further indicated by the frequency of preferences for specific approaches by those who prefer a "without central city" solution. In both areas the percentage point difference by tax attitude is largest in the township alone approach. The highest frequency of preference for this solution is reported by fringe dwellers who respond unfavorably to a tax increase.

In view of the finding that the frequency differential in favor of higher taxes is confined to residents who also prefer solutions to fringe problems in which the central city is an active participant, it becomes pertinent to inquire whether that finding derives from a peculiar distribution of types of household heads. In other words, is the observed association of attitudes uniform in both central city and fringe populations, or is it found only in certain kinds of people? The two populations are, in fact, differently constituted. Household heads residing in the central city are older, more highly educated, have smaller sizes of households, and have fewer complete family units including children of school age than the household heads occuping the fringe. Moreover, in an analysis reported elsewhere it was observed that preferences for solutions involving the central city occurred most frequently among household heads under 45 years of age, with 10 years or more of education, with high weekly earnings and with large families which include children of school age.[4]

Before turning to that phase of the analysis, however, it is well to observe the distribution of attitudes by demographic characteristics. In general, the willingness to accept higher taxes is most frequent among young household heads with large households including children under 17 years of age, and who are in the higher social-economic status cate-

[4] Basil B. Zimmer and Amos H. Hawley, "Approaches to the Solution of Fringe Problems: Preferences of Residents in the Flint Metropolitan Area." *Public Administration Review*, Winter 1956.

gories. The one exception to this general observation occurs in the attitudes of fringe residents sorted by income. It is also apparent that a preponderance of favorable attitudes toward increased taxes obtains in every category of the fringe population, whereas in the central city a clear majority favoring a tax increase occurs only among heads less than 45 years of age and heads with 4 or more persons in their households.[5] proportion favoring as compared to opposed to increased property taxes who also prefer solutions which involve central city participation in all demographic classes of the central city population. It is to be noted, however, that the percentage point difference as between favorable and unfavorable attitudes is greatest among household heads who are under 45 years of age, have least education, earn less than 100 dollars per week, are engaged in manual occupations, have small households and have school-age children. Household heads with these characteristics are more inclined toward solutions of fringe problems which do not require central city participation, especially in cases in which an unfavorable attitude toward tax increase is held.

TABLE IV—PERCENT OF POPULATION WITH GIVEN ATTITUDES TOWARD INCREASED PROPERTY TAXES WHO PREFER SOLUTIONS *With Central City*, BY PLACE OF RESIDENCE AND CHARACTERISTICS OF HEADS OF HOUSEHOLDS

Characteristics of Heads of Households	Attitude Toward Increased Property Tax			
	Central City		Fringe	
	Favorable	Unfavorable	Favorable	Unfavorable
All Household Heads	79.2	63.3	43.2	35.4
Age				
Under 45	76.7	54.7	49.0	37.5
45 and over	82.2	67.6	35.4	34.8
Education				
10 yrs. or less	71.8	61.6	38.9	34.0
More than 10 yrs.	83.6	64.1	47.5	39.3
Occupation				
Manual	76.3	59.1	43.6	35.0
White Collar	88.3	67.5	43.8	30.0
Weekly Income				
Under $100	75.0	60.5	42.7	28.9
$100 and over	88.9	63.4	45.0	50.0
Size of Household				
Less than 4 persons	77.2	64.4	37.2	40.9
4 persons or more	80.7	61.1	47.3	28.6
Type of Household				
With children under 17 yrs. of age	77.6	56.5	47.2	31.1
Without children under 17 yrs. of age	81.4	69.0	35.9	41.2

Taking up now the application of demographic characteristics as controls, Table IV shows that there is a consistently greater

[5] For a more detailed treatment of this subject see the authors' "Property Related Characteristics of Home Owners and Attitude Toward Tax Increase, City and Fringe—Flint Metropolitan Area," *The Journal of The American Institute of Planners*, Spring 1956.

But in the fringe a somewhat different situation is encountered. The convergence of receptiveness to higher property taxes and preference for "with central city" solutions is pronounced among household heads who are under 45 years of age, have 10 years or more education, earn a weekly spendable income of less than $100.00, have households of 4 or more persons, and have children of school age. Occupational differences seem to have no bearing on the relationship under investigation. In several instances the relationship is reversed unexpectedly; that is, a larger proportion of fringe residents who are opposed to property tax increases prefer solutions involving the central city than occurs among those who favor higher taxes. This occurs among household heads whose incomes exceed $100.00 per week, whose households include less than four members, and who have no children under 17 years of age. It is not possible with the data available to determine whether these respondents believe that solutions to fringe problems in which the central city is a participant can be developed without an increase in taxes or whether their unfavorable responses to the suggestion of a property tax increase were based on an assumed maintenance of the status quo of administrative organization. There is some evidence, admittedly inconclusive, that persons in the upper income group and who have small households with no children of school age are among the least convinced that the central city provides as much service for the tax dollar as do township and country governments in the fringe.

A second set of variables which may have some bearing on the relationship in question concerns the kinds of residential experience represented in the population of household heads. Three such variables are used as controls in Table V. In the central city population the highest frequency of respondents preferring "with central city" solutions occurs among those who hold favorable attitudes toward a tax increase. But of greater interest is the fact that the percentage point difference between favorable and unfavorable attitude groups narrows from farm, to village, to city. Evidently the differential association of tax attitudes with type of preferred solution to fringe problems diminishes with the degree to which previous residential experience has been urban in character. This does not apply, however, to the fringe population. A substantial percentage point difference is found only among those heads of households whose last previous place of residence was "village." Years of residence in the locale, i.e., Genesee County, is also a variable of consequence. Central city resi-

TABLE V—Percent of Population with Given Attitudes Toward Increased Property Taxes Who Prefer Solutions *With Central City*, by Place of Residence and Type of Residential Experience

Residential Experience	Attitude Toward Increased Property Tax			
	Central City		Fringe	
	Favorable	Unfavorable	Favorable	Unfavorable
All Household Heads	79.2	63.3	43.2	35.4
Last Place of Residence				
Farm	73.3	55.6	42.6	41.6
Village	73.9	57.9	45.2	16.7
City	50.0	68.9	44.8	42.3
Years of Residence in County				
Under 10 yrs.	66.6	50.1	58.5	25.0
10 yrs. and over	79.4	64.3	38.3	38.7
Local Residential Change				
Never lived in opposite place	87.5	63.8	55.0	44.5
Lived in opposite place	47.6	61.1	40.0	32.8

dents who have lived 10 years or more in Genesee County have the highest perference for "with central city" solutions in both favorable and unfavorable tax attitude groups. But the percentage point differences are approximately the same in the two years of residence categories. In the fringe population, on the other hand, there is a striking difference in the less than 10 years of residence class: 58 percent of those with favorable tax attitudes as against 25 percent of the unfavorable attitude group prefer "with central city" solutions. No difference appears in the over 10 years of residence class of fringe residents. A third type of residential experience results from having or not having moved from one part of the metropolitan area to the other.[6] When this variable is held constant it is found that heads of households who have never lived in the opposite place and who have favorable attitudes toward higher taxes show the greatest frequencies of preferences for "with central city" solutions. The percentage point differences are also largest in that group. But having lived in the opposite place reduces the preferences for solutions in which the central city is a participant, especially among those who hold favorable attitudes toward tax increase in the central city.

As suggested earlier, it is also possible that the relationship of attitude toward property tax increase and preferred solution of fringe problems may be contingent in some way on the characteristics of residential sites occupied by respondents. The data provide some support for this expectation. In the central city favorable attitudes toward an increased property tax are most frequent among household heads whose residential lots are large, whereas in the fringe the occupants of small lots are most favorable. In both central city and fringe, residents whose properties are served by gravel or dirt streets show the greatest willingness to submit to higher taxes. Police protection as represented by frequency of patrolling is, in the data reported, a matter of the respondent observation and judgment rather than of fact. Hence, the differences that appear in the central city population may only reflect errors of observation or a "difference" in standards of regularity. Interestingly enough, the fringe residents show no differences in this respect. In the fringe, reliance on volunteer firemen for fire protection is associated with a greater willingness to pay higher taxes than in the case where fire departments have some full-time firemen. Type of water supply, however, seems to have no differentiating effect on attitudes toward a tax increase.

When the several site characteristics are employed as controls the results are as shown in Table VI. Size of lot is observed to have a greater effect on the relationship of willingness to pay higher taxes to preference for fringe solutions involving the central city in the core city than in the outlying area. That is, in the central city 82 percent of the small lot occupants who favor higher taxes prefer a solution in which the central city is a participant as against 61 percent of those who oppose a tax increase. Among the large lot occupants in the central city, a smaller percentage of those with favorable than of those with unfavorable attitudes toward increased taxes prefer "with central city" solutions. No appreciable difference by size of lot is observable among fringe residents. The pattern of rates found when street surfacing is used as a control on the attitudes of central city household heads is much the same as that observed with size of lot as a control. It is not so in the fringe, however. Among the fringe residents living on hard surfaced roads or streets the proportion preferring "with central city" solutions is lower in those with favorable (39.0 percent) than in those with unfavorable (41.7 percent) attitudes toward a tax increase. But in the heads of households situated on unsurfaced streets there is a marked convergence of willingness to pay higher taxes and preferences for solutions involving the central city. Type of police protection is associated with the relationship in much the same way as street surfacing in both central city and fringe.

Since the type of fire protection and the type of water supply are standard over the entire central city, no opportunity for the observation of differential "effects" is present. Various kinds of facilities for both fire protection and water supply are found in the fringe. It is to be noted, in Table VI, that type of fire protection seems to have no "effect" on the association of willingness to pay higher taxes and preference for solutions

[6] For an extensive examination of this variable as related to type of solution to fringe problem preferred see Zimmer and Hawley, "Approaches to the Solution of Fringe Problems: Preferences of Residents in the Flint Metropolitan Area," *Public Administration Review*, Winter 1956.

TABLE VI—PERCENT OF POPULATION WITH GIVEN ATTITUDES TOWARD INCREASED PROPERTY TAXES WHO PREFER SOLUTIONS *With Central City*, BY PLACE OF RESIDENCE AND SELECTED RESIDENTIAL SITE CHARACTERISTICS

Site Characteristics	Attitude Toward Increased Property Tax			
	Central City		Fringe	
	Favorable	Unfavorable	Favorable	Unfavorable
All Household Heads	79.2	63.3	43.2	35.4
Size of Lot				
Small[1]	82.2	60.7	44.9	38.5
Large	69.2	73.1	43.8	42.9
Street Surfacing				
Hard surface	81.2	60.7	39.0	41.7
Other	66.7	78.9	44.7	33.9
Police Protection				
Regularly patrolled	80.2	63.2	44.0	43.3
Not regularly patrolled	63.6	64.0	40.5	34.2
Fire Protection				
Some full time firemen	.....	.....	43.3	33.3
Volunteer only	.....	.....	45.7	36.4
Water Supply				
Public system	.....	.....	41.8	41.7
Other	.....	.....	44.1	33.9

[1] In central city 60 feet frontage or less; in fringe 75 feet frontage or less.

involving the central city. Whether fringe residents have fire departments staffed by some full-time workers or by volunteers only there is a substantially higher incidence of preference for "with central city" solutions among those willing to pay higher taxes than among those opposed to higher taxes. A differential appears, however, when type of water supply is controlled. Fringe residents served by a public water system are uniformly ready to support solutions involving the central city whether their attitudes toward increased taxes are favorable or unfavorable. But among those who do not have access to a public water supply there is a considerably larger proportion of preferences for "with central city" solutions among residents who are willing to pay higher taxes.

In summary, our data show that there is a significant difference in the type of solution to fringe problems preferred by central city and fringe residents. Central city residents are more likely to prefer a joint city-fringe approach whereas a majority of the fringe residents prefer a solution that excludes a relationship with the central. Nonetheless, even in the fringe more than two-fifths of the household heads prefer a joint solution.

Fringe dwellers as compared with city dwellers have a much larger proportion of household heads who view a tax increase favorably. Although the two areas differ markedly in the amount of property tax paid during the preceding tax year the overrepresentation of favorable attitudes toward a tax increase in the fringe is found to be independent of this difference and is presumed to reflect discrepancies in the level of urban services provided in the two areas.[7]

[7] In an earlier study reported elsewhere it was found that willingness to pay more in taxes was inversely related to level of satisfaction with the services provided in both areas. Zimmer and Hawley, "Property Related Characteristics of Home Owners and Attitude Toward Tax Increase, City and Fringe-Flint Metropolitan Area," *The Journal of the American Institute of Planners*, Spring 1956.

Among both central city and fringe dwellers a relationship is observed between type of solution preferred and attitude toward a tax increase. In each major section of the metropolitan area household heads who responded favorably to a tax increase are more likely to prefer a joint city-fringe approach to the solution of fringe problems than are those who reported that they were opposed to any tax increase. Conversely, those who expressed an unwillingness to pay more in taxes are more likely to prefer a solution that does not involve the central city. It is, however, further observed that the relationship between tax attitudes and preferred solutions is influenced by several demographic, residential experience and site variables.

It appears that in the central city the association of a favorable attitude toward higher property taxes and preference for solutions to fringe problems which involve the central city is a function of age, income, white collar employment, and of never having lived in the fringe. In the fringe that association is most prominent among young household heads in the higher social-economic classes who have complete families which include school-age children and, in addition, who have lived in the area less than 10 years but have never lived in the central city. Presumably any significant changes in the composition of either population would alter the relationships of attitudes. Moreover, any substantial interchange of population between the central city and fringe would probably reduce the proportions of those with favorable attitudes toward tax increases who prefer "with central city" solutions to fringe problems.

It is concluded that the importance usually attributed to this variable as an obstacle to setting up a more inclusive form of government to handle area wide problems has perhaps received more attention than it merits. Our data indicate that a reluctance to participate in a joint effort with the core city because of a fear of higher taxes pertains only to a small proportion of select fringe dwellers. A substantial majority of the household heads in the fringe are willing to pay more in taxes and such persons disproportionately prefer at least a cooperative effort.

There is, of course, no way of knowing *a priori* to what extent these findings can be generalized to other metropolitan areas. The Flint area is somewhat unique in the degree to which the populations living within and outside of the central city are similar in socio-economic and other demographic characteristics. Other than that, however, the Flint metropolitan area displays no prominent peculiarities. The problems confronting it are not unlike those facing other moderately sized metropolitan areas in either their kind or their severity. It seems probable that results corresponding to those obtained here would be discovered elsewhere. Confirmation of this will have to await parallel studies in other metropolitan areas.

BASIL G. ZIMMER
AMOS H. HAWLEY
Department of Sociology, University of Michigan

Methods of Estimating the Postcensal Population of Census Tracts

This study tests the accuracy of selected methods of estimating the current population of census tracts. Eight series of 1950 estimates based on well-known techniques were computed for the twenty-nine Honolulu census tracts. These estimates were then compared with 1950 U. S. Census data to gauge the relative accuracy of each method.

The following techniques were included in the tests: arithmetic, geometric, and ratio projection, two forms of the dwelling capacity method, proration, the censal ratio method, and a combination of the latter two techniques. All estimates referred to residential (rather than "daytime") population.

Arithmetic and Geometric Projection

Arithmetic and geometric projection are based on the assumption that trends in population growth observed during the most recent intercensal period will continue unchanged during the postcensal period. In arithmetic projection, the intercensal *amount* of increase (or decrease) is projected; in

geometric projection, it is the *percent* of change that is projected. Both methods have been in common use for many years.

Tests were made by projecting 1930-1940 trends to 1950 and comparing the resulting figures with the corresponding census counts. (Although census tracts were not defined for Honolulu prior to the 1940 Census, it was possible to make accurate estimates for 1930 from unpublished data for enumeration districts.) The projected figures were not adjusted to sum to an independent citywide estimate, as was the case for other methods.

The 1950 estimates differed considerably from the corresponding census counts, both for arithmetic and geometric techniques. Average error was 25.3 percent for the former and 25.6 percent for the latter. Median error was respectively 22.0 percent and 15.6 percent. Error was greatest for census tract 12 (the central business district), amounting to 107.2 percent for arithmetic projection and 112.6 percent for the geometric method. Tract 1, an area of military housing near Pearl Harbor, ranked second in percent error, with 69.3 percent for arithmetic and 69.2 percent for geometric projection.

Ratio Method

In the ratio method past trends in tract-to-city ratios are projected beyond the most recent census, and applied to an independent current estimate for the city as a whole. Ratios can be projected arithmetically (as in the present study), geometrically, logistically, or perhaps in some other fashion. Census tract 26, for example, had 3.5 percent of the citywide total in 1930 and 3.9 percent in 1940; the estimated ratio for 1950, projected from these percentages, was 4.3 percent. This ratio was then applied to an absolute figure estimated independently for the city as a whole. (The actual 1950 U. S. Census total for Honolulu was used in the test, but an estimated total would have to be used in actual practice.[1])

The ratio method proved little different from arithmetic or geometric projection in estimating accuracy. Average error for this technique was 27.9 percent, highest for any method tested, and median error was 21.5 percent, second highest of all eight techniques. As in the case of arithmetic and geometric projection, errors were greatest for tract 12 (with 124.2 percent) and tract 1 (74.7 percent).

Dwelling Capacity Method

Population estimates made by the dwelling capacity method are derived from corresponding estimates for total dwelling units and population per unit. The number of dwelling units in a given census tract is, of course, the sum of the number reported by the most recent U. S. Census of Housing plus the approximate number constructed in that tract since the Census date. Postcensal construction is based in turn on building permit data or utilities connections. Wherever possible, statistics on conversions and demolitions should also be considered. The current dwelling inventory is then multiplied by the estimated population per unit, by tract, which is either taken directly from the most recent Census or extrapolated from data for two previous Censuses. Results are then adjusted to sum to an independent citywide population estimate.

Two variations on the dwelling capacity method were tested. In the first, the number of units built since 1940 was taken from the 1950 Census of Housing (in actual practice, of course, it would be necessary to use building permit or utility data) and added to the number reported in the 1940 Census without allowance for conversions or demolitions. In the second test, the 1950 dwelling unit inventory was transcribed directly from the 1950 Census totals, thus assuming perfect knowledge of "postcensal" (1940-1950) residential construction, conversion, and demolition for each tract. (Here, too, actual application of the method would require use of building permit or utilities data, thus introducing the possibility of errors not encountered in the present tests. Recourse to Census data for the tests was dictated by the impracticality of reviewing Honolulu building permit data for the ten-year "postcensal" period under study.) In both variations of the dwelling capacity method, population per dwelling unit for 1950 was assumed to be the same by tract as in 1940. The 1950 population estimates thus computed for each tract were then ad-

[1] For recommended techniques, see: Jacob S. Siegel, Henry S. Shryock, Jr., and Benjamin Greenberg, "Accuracy of Postcensal Estimates of Population for States and Cities," *American Sociological Review*, August 1954, pp. 440-446; Robert C. Schmitt, "A Method of Estimating the Population of Cities," *American Journal of Public Health*, November 1954, pp. 1426-1427; U. S. Bureau of the Census, "Illustrative Example of a Method of Estimating the Current Population of Subdivisions of the United States," *Current Population Reports*, Series P-25, No. 133, March 16, 1956.

TABLE I—PERCENT ERROR IN 1950 POPULATION ESTIMATES FOR HONOLULU CENSUS TRACTS

Method	Number of census tracts, by percent error						Average error (percent)	Median error (percent)
	All tracts	Under 5.0	5.0 to 9.9	10.0 to 24.9	25.0 to 49.9	50.0 and over		
Projection methods								
Arithmetic	29	5	4	9	8	3	25.3	22.0
Geometric	29	6	3	9	6	5	25.6	15.6
Ratio	29	5	5	8	6	5	27.9	21.5
Dwelling capacity method								
Demolitions disregarded	29	7	9	8	3	2	14.2	8.3
Demolitions considered	29	8	8	10	2	1	12.8	9.2
Vital rates methods								
Proration	29	12	5	9	3	..	10.6	5.9
Censal ratio	29	9	5	9	5	1	15.4	10.0
Combined	29	8	10	8	3	..	10.9	7.1

justed to sum to an "independent city estimate," actually the enumerated total taken from the 1950 Census.

Accuracy was far superior to that found for the three projection methods. Average error was 14.2 percent when demolitions and conversions were disregarded and 12.8 percent when they were taken into account. Median error was respectively 8.3 percent and 9.2 percent. Tract 1, the area containing a good deal of military housing, once again made the poorest showing, with errors of 92.1 percent (demolitions disregarded) and 91.9 percent (demolitions considered). This degree of error resulted chiefly from postcensal changes in population per dwelling unit, a ratio which in 1940 (but not 1950) was distorted by the large proportion of residents in nondwelling-unit quarters. Tract 12, the central business district, fared second worst (with an error of 58.7 percent) when demolitions were disregarded but had the least relative error when demolitions were considered. The second highest error for the "demolitions considered" series was 30.7 percent, found for tract 9 on the downtown fringe.

The dwelling capacity method is apparently the only one subjected to systematic testing at the census tract level in recent years. This technique was used to make 1950 estimates for Seattle census tracts, using actual building permits (unadjusted for demolitions) to determine postcensal increases in the number of dwelling units.[2] The resulting population estimates, when compared to preliminary tabulations of the 1950 U. S. Census, were found to have an average error of only 8.5 percent.

Proration

In the proration method the estimated population is distributed by census tract in proportion to the tract-by-tract distribution of such symptomatic data as live births, deaths (both on a residence basis), active utility connections, city directory entries, school enrollment, mail boxes, or registered voters. A census tract with 4.7 percent of the city's residential water meters (for example) is assumed to have a corresponding share of the city's population. This method thus requires an independent estimate of total population for the city. Unlike the other techniques included in the present comparison, it does not depend on data taken from one or more decennial Censuses.

In the present test, civilian population was allocated on the basis of total civilian vital events, that is, the sum of live births and deaths, available on a place of residence

[2] Robert C. Schmitt, "Estimating Current Populations of Census Tracts," *Sociology and Social Research*, September-October 1952, pp. 12-15.

basis from the Territorial Department of Health. Lack of data made it necessary to confine the analysis to civilian residents but the principles and sources described above would probably apply equally well to total populations. As with other methods the actual 1950 Census figure for Honolulu was substituted for the independent citywide estimate required in actual practice.

Accuracy of the proration method surpassed that found for any other technique tested. Average error was only 10.6 percent, and the median was only 5.9 percent. Error was greatest in the case of tract 1, the military area near Pearl Harbor, with 48.5 percent, and tract 5, Iwilei industrial district, with 30.4 percent. The method proved especially good for residential areas.

Censal Ratio Method

The censal ratio method derives postcensal trends in population from corresponding changes in symptomatic data. It involves three steps: 1) computation of census-year ratios of population to births, deaths, water meters, school enrollment, or similar symptomatic data for each census tract; 2) application of census-year ratios to postcensal symptomatic data to obtain tentative postcensal population estimates by tract; 3) adjustment of the tentative estimates to make them total an independent citywide estimate.

The test of the censal ratio method differed from tests for other techniques in that an eight-year (rather than ten-year) postcensal period was used. This modification was necessary because of the unavailability of symptomatic data for years prior to 1941, and the resulting need to base census-year ratios on data from the OCD census taken March 29, 1942 rather than on the 1940 decennial count. Lack of data also made it necessary to confine the estimates to civilian population rather than the total, civilian and military combined. As in the test of the proration method, civilian vital events (the sum of births and deaths) were chosen as symptomatic series and the actual 1950 civilian population of Honolulu was used as a control total.

The censal ratio method proved less successful than either proration or the dwelling capacity method in estimating the population of census tracts. Average error amounted to 15.4 percent, and median error was 10.0 percent. Results were poorest for tract 22 (Waikiki, an apartment and tourist hotel area), with an error of 91.1 percent, and tract 21 (Manoa Valley, a high-rent area invaded by public housing since 1940), with 47.3 percent. The estimating error for tract 1, however, was reduced to 17.5 percent. The inferiority of the censal ratio method to proration in this test suggests that the ratio of population to vital events generally varies less from tract to tract in a given year than from year to year for a given tract.

Combined Method

A final test involved combination of the proration and censal ratio methods. Estimates made by each method, as described above, were added and averaged, in the hope that an overestimate resulting from one technique would tend to cancel an underestimate produced by the other.

This combination of methods gave results somewhat inferior to the better of the two. Average error was 10.9 percent, slightly more than that found for proration. Median error was 7.1 percent. As in the case of the censal ratio method, error was greatest for tracts 22 and 21 (respectively 45.8 and 34.3 percent). Tract 1, which proved most difficult to estimate accurately by six of the methods (or methodological variations) tested, fared much better, with an error of only 15.5 percent resulting from the combined method.

Summary

Although the present test can hardly be regarded as conclusive—it reports data for only eight of many possible combinations of method and symptomatic series, for a single city and a single year—it nevertheless suggests the following tentative conclusions: (1) At the census tract level, postcensal population estimates based on projection of past trends are vastly inferior to those derived from current symptomatic data (vital events in the present instance). (2) Among techniques using symptomatic data, proration appears to be both the simplest and most promising. (3) No matter what method is used, results will be much more accurate for residential areas than for industrial, tourist, military, and central retail and fringe areas. (4) Even with the most accurate techniques, relatively high average errors and extremely large errors in estimates for individual tracts are to be expected.

ROBERT C. SCHMITT

Redevelopment Research Analyst,
Honolulu Redevelopment Agency

Measuring the Suburbanization of Manufacture

STATISTICAL studies of the process variously described as decentralization, deconcentration, diffusion, or suburbanization of manufacturing and other economic activities have failed to confirm the impression of rapid locational shifts since the prewar period. The data seem to belie the conclusions one may draw from observing the thousands of new factories and other establishments built during and after the war in the outskirts of our metropolitan areas. The purpose of this note is to help resolve the apparent conflict between statistical and visual observation.

For this purpose we shall accept Donald J. Bogue's term "suburbanization" as denoting an increase in the proportion of the total manufacturing or other activity of a standard metropolitan area that takes place outside the central city. For the sake of simplicity the discussion will be limited to manufacture. The latest statistical study along this line, and one of massive scale, was prepared by Kitagawa and Bogue.[1] The results seem to establish beyond doubt that, for aggregates of metropolitan or industrial areas, there was little or no suburbanization of manufacturing between 1939 and 1947, the latest year for which data were available. This inclusion is drawn from the fact that the percentage distribution as between city and periphery of manufacturing establishments, production workers, and value added by manufacture shows very small changes. There is a real question, however, whether these statistical findings—correct as far as they go in indicating the locational distribution of *activity*—really deny the presence of substantial shifts in the location of manufacturing *facilities*. While the authors do not jump to this possibly unwarranted conclusion, many users of their materials do, and references in the study to an apparent conflict between common observation and statistical "fact" are conducive to misinterpretation.

The key to resolving this conflict lies in the rate of utilization of manufacturing facilities. In 1939, a large proportion of manufacturing facilities throughout the country was still idle, being either completely vacant or used at less than capacity. The proportion of idle facilities was typically much larger in the central cities than in the outlying areas. This observation is supported by the data used in the study. In the central cities of principal standard metropolitan areas, the number of production workers in manufacturing industries fell from about 4.3 million in 1929 to 3.5 million in 1939, or by 18 percent. In the "rings" of these areas as defined by Kitagawa and Bogue, the decline was from 2.058 million tc 2.011 million, or only 2 percent. Value added by manufacture likewise dropped by 32 percent in the central cities but by only 11 percent in the "ring" areas.[2]

As the demand for manufactured products increased during and after the war, the idle facilities that were concentrated in the central cities were pressed into service with or without modernization; and employment and value added in these facilities soared even though there may have been little expansion in centrally-located capacity. Consequently, the locational distribution of *activity* between central city and periphery, as measured by employment and value added, was about the same in 1947 as in 1939. However, if a measure of the volume of *physical facilities* such as floor space of usable area was available, it may well show a vastly greater increase in peripheral facilities over 1939 than in central facilities.

Equal rates of increase in utilized capacity in central and peripheral areas (which is essentially what the study shows) are entirely consistent with highly differentiated rates of increase in *total* capacity, provided there is a substantial initial difference in capacity utilization as between the two types of areas. This is outlined in the following schematic illustration:

	1939	1947
Central city		
Utilized 1939 capacity	60	100
Additional capacity 1939-1947	....	20
TOTAL UTILIZED	60	120
Periphery		
Utilized 1939 capacity	90	100
Additional capacity 1939-1947	....	80
TOTAL UTILIZED	90	180

[1] Evelyn M. Kitagawa and Donald J. Bogue, *Suburbanization of Manufacturing Activity within Standard Metropolitan Areas* (published jointly by Scripps Foundation for Research in Population Problems, Miami University, and Population Research and Training Center, University of Chicago, 1955).

[2] Kitagawa and Bogue, *op. cit.*, Table II-2.

In the example, the total utilized capacity is doubled from 1939 to 1947 in both the central city and the periphery, and whatever the 1939 ratio of utilized capacity in the periphery to utilized capacity in the central city may have been, it would be found to be the same in 1947. However, this finding obscures an equally or more important fact: the capacity added in the peripheral area between 1939 and 1947 was four times as large as the capacity added in the central city. This, of course, is the kind of visual observation which is seemingly (but not really) contradicted by the battery of statistics.

The presumption that locational differences in the rate of utilization of manufacturing facilities between 1939 and 1947 may have produced an illusory picture of persistence in location patterns is strengthened by a comparison of activity data for 1929 and 1947, both of which periods can be considered years of about full utilization of capacity. Between 1929 and 1947 the number of production workers in manufacturing industries rose by about 50 percent in the "ring" areas and by less than 25 percent in the central cities. Value added increased by about 150 percent in the "ring" areas and by little over 100 percent in the central cities. The ring areas accounted for roughly 37 percent of the production workers in 1947 as against 33 percent in 1929, and for 36 percent of value added in 1947 as against 30 percent in 1929.[3] Thus, taking the period from 1929 to 1947 as a whole, there seems to have been a movement toward suburbanization.

The importance of various degrees of capacity utilization to changes in the locational distribution of activities was fully recognized in the earlier study of industrial decentralization by Coleman Woodbury.[4] It is also stressed by the local planning practitioners who contributed non-statistical reports to the Kitagawa-Bogue work. Several of these reports, in trying to explain low rates of suburbanization, emphasize the resurgence of manufacturing activity in previously idle central facilities. It is all the more remarkable that the statistical analysts have made so little of this point. There are, of course, no adequate measures of physical facilities they could have used as substitutes for, or supplements to, the activity data, but the absence of such measures clearly limits the conclusions suggested by the activity data.

When detailed figures from the 1954 Census of Manufactures become available for comparison with 1947, they may well afford better insight into changes in the locational pattern of manufacturing and other facilities, even though they are limited to evidence of "activity" such as employment and value added. For the 1954 data, as well as those for 1947, would cover periods of fairly full capacity utilization. Should they show a movement toward suburbanization after 1947, however, it would yet be hazardous to consider this to be a break with the 1939-1947 experience which was so greatly influenced by locational differences in the rate of utilization. Moreover, it will still be important to the analysis of the 1947-1954 changes in *individual* metropolitan areas to take account of local shifts in capacity utilization.

If the observations in this note are valid, they emphasize the importance of learning more about locational changes in the *inventory* of facilities, whether used or unused, and of guarding against conclusions drawn solely from activity data. Planners and other local officials, as well as students of urban growth, are at least as much interested in locational changes in the distribution of facilities as they are in locational patterns of activities. While the compilation of national inventories of physical facilities for manufacturing or other economic activities is probably beyond reach, local inventory data, vital for both analytical and planning purposes, are often less difficult to obtain. The case at hand highlights also the usefulness to urban land use studies of the concept of real estate inventories and rates of inventory utilization.

LEO GREBLER

Council of Economic Advisers,
Washington, D. C.

[3] *Op. cit.*, Table II-2.

[4] Coleman Woodbury with the assistance of Frank Cliffe, "Industrial Location and Urban Redevelopment," Part II, in Coleman Woodbury (ed.), *Future of Cities and Urban Redevelopment* (Chicago: University of Chicago Press, 1953), Chapter V.

Fallacious Statistical Estimation Based upon Regression Analysis

IN connection with the rate case brought before the California Public Utilities Commission by the Pacific Telephone and Telegraph Company in 1948 (Application No. 28211; Decision No. 41416), the Commission Staff prepared a *Report on Determination of the Reasonable Amount of American Telephone and Telegraph Company's Service Costs Allocable to Operating Expenses of the Pacific Telephone and Telegraph System State of California for the Year 1946.* This report was designated Special Study No. S-357. It is the purpose of the present paper to criticize a certain statistical estimate made therein. I should, perhaps, not bother to do so if the report in question were a private paper; or if the statistical estimate in point had been entered as evidence in litigation between private parties. But since it was entered as a statistical argument in connection with an issue of public policy, I believe it deserves public scrutiny. What is more, if allowed to stand as an unchallenged precedent, this type of statistical argument might be adopted in other regulatory investigations with equally questionable or misleading results.[1]

The figure in question is embodied in the following quotation ". . . . It is estimated that the expenses (excluding taxes) of the American Telephone and Telegraph Company, incurred in its interest as an investor in the year 1946, were not less than 0.15% of its 1946 average investment in subsidiary and affiliated companies."[2] The analysis which follows, and the opinion which results, is confined solely to the statistical aspects of the estimating procedure which the Commission Staff evidently employed.

In S-357, itself, there is no complete and explicit statement of the statistical means by which the estimate of 0.15% was made. The nearest I find there to a description of the actual estimating procedure is the following:

"To obtain a yardstick with which to measure the approximate magnitude of the expenses incurred by reason of investments in affiliated companies, there have been developed ratios of the expenses, exclusive of taxes, incurred by General Telephone Corporation and 12 registered public utility holding companies to the average investments which these companies have in the securities of subsidiary and other companies. The ratios thus developed afford a measure of the costs incurred by a company in administering and protecting its investments in affiliated companies, in the interest of its investors, as contrasted with costs incurred in furnishing services to operating companies of the system.

"The tabulation [see page 383] show the ratios of expenses to average investments for the years 1944, 1945, and 1946, by companies.

"Based upon the study of the costs of American Telephone and Telegraph Company functions identifiable in whole or in part with its interest as an investor, and upon the above ratios of expense to investments, it is estimated that the expenses (excluding taxes) of the American Telephone and Telegraph Company, incurred in its interest as an investor in the year 1946, were not less than 0.15% of its 1946 average investment in subsidiary and affiliated companies."[3]

Further enlightenment can, however, be found in the testimony of a Research Engineer on the Commission Staff, given under cross-examination before the Commission. The following interchange (transcript 4946) took place:

Q. Yet, if you drew a curve through plotted material wouldn't the ratio for a company having an investment of $3,258,000,000 have dropped away down on the line, maybe to .03 percent somewhere along in that zone?

A. No, sir.

Q. Rather than 0.15%?

A. No, sir, I don't think so. I did draw such a curve and the curve flattened out materially, and while it is quite an extrapolation to go the size of the American Company, I would have preferred to have companies more nearly the size of the American Company, but there were none, it appeared from the shape of that curve out at the largest investment, Electric Bond and Share, that the curve would not drop greatly below that level.

[1] Even so, I might not bother to enter upon this criticism if the estimate in question had perished with the 1948 rate case. But in 1954 the Commission Staff was still using the very same estimate in hearings associated with a rate case then current. This will become apparent presently.

[2] *Special Study* No. S-357, p. 5.

[3] *Special Study* No. S-357, pp. 4-5.

Company	Average Investments Year 1946	Ratio—Expenses Excl. Taxes to Average Investments		
		1946	1945	1944
General Telephone Corp.	$ 34,990,000	0.758%	0.402%	0.377%
American Gas & Elec. Co.	134,514,000	.233	.209	.195
American Light & Traction Co.	112,095,000	.284	.362	.253
American Power & Light Co.	257,303,000	.183	.197	.186
Columbia Gas & Elec. Corp.	248,639,000	.398	.380	.477
Commonwealth & Southern Corp.	306,630,000	.204	.170	.158
Consolidated Elec. & Gas Co.	34,874,000	.341	.317	.360
Electric Bond & Share Co.	447,451,000	.218	.199	.172
Electric Power & Light Corp.	151,819,000	.557	.334	.374
Middle West Corp.	67,249,000	.419	.472	.397
New England Public Service Corp.	29,414,000	.510	.404	.265
North American Company	203,166,000	.582	.455	.383
Standard Gas & Elec. Co.	165,337,000	.350	.286	.235

(*Source:* Moody's *Manual of Public Utilities*—1947)

As mentioned earlier, the same estimate was being used by the Commission Staff as recently as 1954. I will not discuss the rather obvious hazards arising from this extrapolation over time, but I cite from Application No. 33935, transcript 5359—5362 to establish that such extrapolation was made. A member of the Commission Staff was under cross-examination on March 18, 1954:

Q. That 9½ percent factor comes, does it not, from the .15 percent factor which was computed in certain studies made in 1946 and 1947?

A. Yes, sir, that is correct.

Q. How was that point .15 percent factor computed?

A. Well, I have generally stated it on page 16 of Exhibit 113 in paragraph 3.

As I understand it, it is a judgments percentage of average investment based on the ratios of General Telephone Corporation and 12 gas and electric holding companies.

Q. That judgment was based upon the ratios?

A. Of the administrative and general expenses excluding taxes to the average investor.

Q. And those ratios were computed on the expenses and average investment as of the time of the study, 1946 and 1947?

A. I believe there is—excuse me, I have a copy of that exhibit 263. The investments for the year 1946 and the ratio of expenses excluding taxes to the investments are shown for the years 1944, 1945 and 1946.

Q. Now, were the companies used in making that study the same size as the American Telephone and Telegraph Company?

A. No, sir. I believe the largest one was roughly one-eighth the size of the American Telephone and Telegraph Company at that time.

Q. Now, how was the figure, the percentage derived, then for the American Telephone and Telegraph Company?

A. Well, I believe that, considering the relationships of these companies, it was the Staff's judgment that 15 hundredths would be applicable to the investment in the American Company.

Q. You have looked over those studies, have you?

A. Yes, sir.

Q. You are familiar with them?

A. Yes, sir, in general.

Q. Now, what method was used in arriving at the .15 percentage for the American Telephone and Telegraph Company? Can you explain it briefly to the Commission?

A. Well, I think there is a trend established with these 13 companies that indicates some decrease in ratio with increase in investment.

Q. As a matter of fact, a curve was drawn?

A. I believe so, through these 13 points.

Q. Did the curve touch the points?

A. I think it was probably fitted to the general trend of all the points. I mean, they would not necessarily all fall on the same line.

Q. And that curve showed a trend so that the larger the investment the smaller the ratio of expense, is that correct?

A. I don't believe it was a straight line relationship, no.

Q. Well, I didn't ask whether it was a straight line relationship. I think you agreed it was a curve.

A. Yes, sir, up to certain limits it would.

Q. Now, what has happened to the American Telephone and Telegraph Company's average investments since 1946 when this ratio was established?

A. I believe they were approximately doubled.

Q. Did you make this original study, ...?

A. No, sir, I did not.

Q. Were you told by anyone to use the same factor?

A. No, sir, that was my own idea. Excuse me, I will qualify that. I didn't use the same factor of .15.

Q. You did not?

A. In this study, I did not, no sir.

Q. What factor did you use?

A. I used the results obtained from the three previous Staff exhibits.

Q. I see. You used the 9.5 ratio which was based upon the .15?

A. That is correct, yes, sir.

Evidently, then, the Commission Staff used what is known as regression analysis as a basis for the estimate. That is to say, they plotted points representing 1946 expense ratios and corresponding average investments for the thirteen companies noted above, passed a curve through these known values, then estimated what the unknown expense ratio for American Telephone and Telegraph Company might be on the basis of the known figure for its average investment in 1946.

This estimating procedure is often thought to give fairly reliable results for estimates *within the limits of the data actually observed* and providing that the data are homogeneous. Even so, the estimates are usually less good as the limits of the observed data are approached, and they are quite untrustworthy if made much beyond those limits.

A book which is something of a classic in statistical literature can be quoted on this point:

"Unless there are good reasons for supposing that the fitted curve is an accurate respresentation of a theoretical relationship, it is dangerous to assume that a fitted parabola can be used outside the range for which it was ascertained.

"It would be instructive for the student to fit merely a segment of some actual series and note how rapidly the curve calculated from the segment diverged from the observations outside its limits. It has been shown that even within the limits of the fitted observations the fit tends to be worst as the limits are approached. The higher powers of x become of greater and greater effect the more we diverge from the centre of the fitted segment and tend, so to speak, to 'wag the tail' of the curve."[4]

Or to take a work which is prominent in the specialized literature of correlation analysis, of which regression analysis is a part:

"It is particularly to be noted that determination of the line or curve of relationship gives no basis for estimating beyond the limits of the values of the independent variable actually observed. No matter whether a formula has been fitted or not, any attempt to make estimates beyond the range of the original data by "extrapolation,' i e., by extending the curve beyond the range of the observed data, gives a result that is not based on the statistical evidence."[5]

Even in elementary texts addressed to the beginning student the point is made: "It must be emphasized that predicting a Y value from an X value beyond the range of the given observations should be done cautiously since the relationship may be operative only within the limits of the observations."[6]

And again: "And this is as good a place as any to warn the student against too confident use of such trend equations or regression equations . . . . and in this case we need to warn the student against the practice of *extrapolation*. Just as *interpolation* is the finding of a value that lies between values we already know, so extrapolation is the projecting of values beyond the limits of our knowledge. The good statistician is very cautious about extrapolation, which can very easily lead him far astray."[7]

The statistical extrapolation apparently made by the Commission Staff is the most daring I have ever seen in a serious piece of

[4] G. Udny Yule and M. G. Kendall, *An Introduction to the Theory of Statistics*, 12th ed., rev., (London: Charles Griffin and Co., 1940), pp. 325-327.

[5] Mordecai Ezekiel, *Methods of Correlation Analysis*, (New York: John Wiley and Sons, 1930), p. 109.

[6] George Simpson and Fritz Kafka, *Basic Statistics*, (New York: W. W. Norton and Co., 1952), p. 380n.

[7] Albert E. Waugh, *Elements of Statistical Method*, 3rd ed., (New York: McGraw-Hill Book Co., 1952), p. 325.

work. The pertinent investment of American Telephone and Telegraph is more than *seven* times as large as the upper limit of the observed data—we have here an extrapolation based upon a projection 630% beyond the limit of known values relating the expense ratio to average investment in 1946.

Another way of bringing home the enormity of the extrapolation is this. The entire range of investment over which the relationship of expense ratio to average investment is observed is from $34,990,000 to $447,451,000 —or $412,461,000. The average investment of American Telephone and Telegraph lay $2,811,173,000 above even the upper limit of that range. In other words, the expense ratio for American Telephone and Telegraph was estimated on the basis of extending a relationship over a gap 680% as wide as the range over which that relationship was observed to hold at all. Such things just shouldn't be done with statistics.

But if such things *were* to be done, let us consider the statistical ambiguity of the curve of relationship itself. In the testimony quoted above the witness said, "I did draw such a curve and the curve flattened out materially . . . . " He did not, however, there, or in Special Study No. S-357, state the equation for the curve he used or explain how it was fitted to the data. Now this is an important omission. For there are a number of curves which can be drawn through the plotted points representing the data of the thirteen companies, all of which will "flatten out materially." And, obviously, the extent to which any one curve flattens out, and the rate at which it flattens out, will have a great deal to do with an estimate of the expense ratio based upon it—especially when that estimate corresponds to an average investment so far beyond the limits of the actual data.

To illustrate this point, I have fitted mathematically several curves to the same data utilized by the Commission Staff. I shall try to be explicit about what I have done. The procedure used to fit these curves is known as the method of least squares, which Yule and Kendall refer to as being "in almost universal use at the present time."[8] The several curves mathematically fitted to the data on Chart I have the equations given below, with identifying numbers for each:

1. $y = 75.72x^{-0.288}$
2. $y = 0.01e^{3.9254e^{-0.0000000006137x}}$
3. $y = 0.5065e^{-0.000000002132x}$

In each case, y represents expense ratio and x represents average investment of the firm in 1946.

If estimation of the expense ratio for American Telephone and Telegraph were attempted by extrapolation based upon its investment figure of $3,258,624,000, the estimated value would depend upon the particular curve used. With estimated ratios identified by the numbers of the respective curves, these values would be:

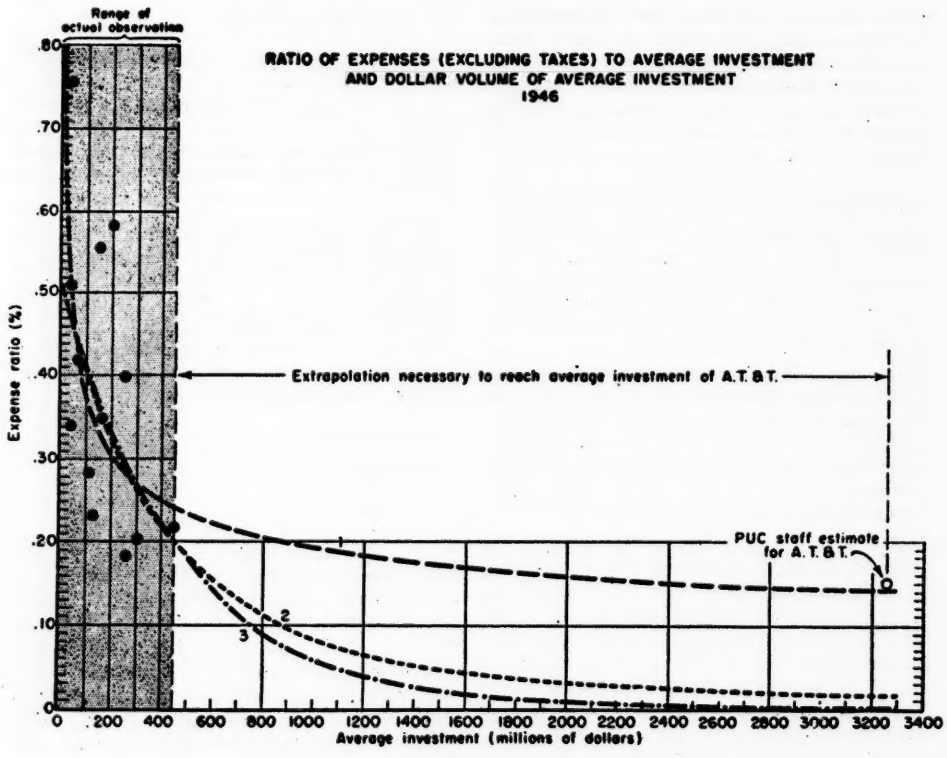
1. 0.138%
2. 0.0017%
3. 0.00049%

It will be noted from Chart I that all of the fitted curves pass through the plotted points of the actual data with a satisfactory representation of whatever relationship appears therein. Yet each gives a different result when extrapolated for attempted estimation; and each gives a lower value than the 0.15% estimated by the Commission Staff. This contrast is particularly striking since the Staff's report states: ". . . . it is estimated that the expenses (excluding taxes) of the American Telephone and Telegraph Company, incurred in its interest as an investor in the year 1946, *were not less than 0.15%* of its 1946 average investment in subsidiary and affiliated companies." And just a little later: "Applying the ratio of 0.15% to this investment figure [$3,258,624,087] yields $4,888,000 *as the minimum amount* of expenses, exclusive of taxes, incurred by the General Department in 1946 for the purpose of administering and protecting these investment."[9] The figure 0.15% is not given as an estimate around which wide latitude is is admitted; it is given as an estimate of the lowest limit to which the true expense ratio for American Telephone and Telegraph could have fallen in 1946.

The fact that all of the curves I have fitted and plotted on Chart I fall below 0.15% as they pass by American Telephone and

[8] *Op. cit.*, p. 311.

[9] Both quotations from *Special Study* No. S-357, p. 5. Italics added.



Telegraph's investment figure of 3,258,624,000 does not mean that the true expense ratio at that point is less than 0.15%. In the earlier part of this paper I eschewed all estimates based upon such gross extrapolations of this type. And I can say here that there are other curves which I could fit mathematically to these same data which would give estimated expense ratios above 0.15%.[10] The point is, there is nothing to choose from among the wide variety of curves that will pass reasonably well through the known data for the thirteen firms which the Commission Staff has examined. Yet, as I have tried to demonstrate, the estimate arrived at by the Staff's technique depends entirely upon the curve which happens to be selected.

In my opinion, the procedure is entirely without validity when carried to the extreme represented here.

ROY W. JASTRAM

Professor of Business Administration,
University of California, Berkeley

[10] One such curve which I did fit would appear to give an estimated expense ratio of 118.9% for AT&T, even though it represented the known data below investment of $447,451,000 sensibly well. Should this be taken to mean that AT&T's expenses incurred in its interest as an investor *exceeded* its average investment in subsidiary and affiliated companies in 1946?

Book Reviews

Governing Urban America. By Charles R. Adrian. New York: McGraw-Hill Book Company, 1955, pp. vi, 452.

This is a textbook designed for collegiate study of modern American city government. It is the best of its kind available today. The author has performed a distinguished service to civic education in providing us with such an excellent tool for preparing young men and women to discharge their local political responsibilities.

The excellence of the volume under consideration stems from a broad and penetrating understanding of the character of the modern urban environment in America, an environment in which most of its prospective readers will surely spend their lives. Not only in chapters dealing with the urbanization of America, the flight to the suburbs, and the future of city government but throughout the work, Adrian brings to his subject a good knowledge of the developing literature of urban sociology, urban psychology, and urban planning and redevelopment. In dealing with politics, pressure groups, and citizen participation, the author also makes effective use of the new insights into the political process gained by modern research into political behavior, group activities, and public attitude analyses. More than most textbook writers, Adrian has treated the old ogres of bossism and municipal corruption in a realistic and "practical" way and while his treatment will not light many fires of crusading zeal, it will not turn youth away from the affairs of city hall with disgust and loathing.

Adrian deserves commendation also for his handling of the legalisms of municipal corporate powers and of the usual strictures on administrative organization, management, and procedure. In most comparable works, these subjects are usually deadly dull and in some cases utterly unrealistic both as to emphasis and as to prescriptions for improvement. By contrast this text gives more attention to urban politics and pressure groups than to administrative organization and municipal law.

This reviewer wishes that the author had provided fuller treatment of municipal services and responsibilities, and particularly that education and recreation had received more space. Obviously, something must be abbreviated in a work of this kind and, on the whole, the services are well-treated in terms of the major contemporary problems involved in providing them in adequate quality and quantity.

The book is generously illustrated with charts, graphs, tables, and cartoons in the best style of modern textbook making.

WILLIAM H. YOUNG

University of Wisconsin

The Metropolis in Modern Life. Edited by Robert Moore Fisher. New York: Doubleday and Company, Inc., 1955, pp. ix and 401.

This book now available on the book market for some time, owes its success probably to the fact that it combines the conventional and the critical. The conventional undoubtedly makes it useful as a text for members of the younger generation who want to find their way into current thought prevailing in the field of the urban sciences; while critical discourses, interspersed in the book, challenge the expert to a revision of previcus findings.

The reader is not confronted with some kind of urban utopia. As chairman of the Bicentennial Celebration at Columbia University, Ernest M. Fisher, the noted land economist, refers to the worth of the book that is inherent in many treatments of urban problems as seen in different parts of the world. There are contributions from many countries. They discuss theorems and conditions which have so far been accepted as

true. They let the reader into the workshop of the urban scientist; revealing however that he may want to find out more about the city rather than accept and distribute knowledge fairly well accepted by his colleagues.

Needless to say, there are many contributions which—as is wont in a symposium of this nature—indicate only too clearly that the author, at the time he made his contribution, was not beset in his work with any problems of general interest. Other contributions are doubtful in their conclusions though always stimulating even though the contribution often has to be rejected by any careful reviewer.

Reading the book at this distance of time, it causes particular amazement with a seemingly complete renunciation of urban life and its peculiar characteristics. Not untypically for the book as a whole, it is pointed out to us that attributes we might have associated with an urban way of life since the days of George Simmel, Robert E. Park, and Louis Wirth, are not necessarily urban at all. We may—asserts the brilliant contribution of Albert C. Reisse, Jr.—have more inventions, more creativity, more of a relativistic spirit: we may have more agricultural pursuits, more division of labor, more social movements, more anonymity and greater size and density of settlement in the city, but all these attributes can also be found outside the city environment. Consequently, we have to revise our idea about typically urban traits.

This statement leads us to two reflections: an historical one and one related to type construction in general. Urban attributes such as those mentioned above are indeed tied closely to historical urban development. Likewise, sporadically urban attributes may also be found in rural areas today, although they did perhaps not occur here to any sizeable extent before the historical advent of industrialization. If the anthropologist tells us that (1) village formation, (2) urbanization, and (3) industrialization are three states of cultural development which occur in all cultures and in the Western world also, then it would have to be proven that urban traits are not simply the outcome of industrialization in the city or the country-side, that do occur in Western Civilization before even industrialization has taken place. It would probably be impossible to undertake such proof in other ways than by using available historical sources, nor does the author of the above-mentioned contribution offer any new empirical proof other than by casual reference to well-known contemporary phenomena.

Type construction, on the other hand, without the opportunity for empirical verification cannot be accepted as a permanent contribution to sociological knowledge or social science knowledge in general. It is perhaps an advantage that this review appears late, and that the reviewer is able to state that no attempt has been made since—not even by the author himself—to prove that urbanism, as so far known and defined, appears in an appreciable number of cases in the non-urban environment. Under the the circumstances we must assume that the author of this contribution once thought very challenging has not offered a real challenge at all. He may, under existing circumstances, have pointed only to the occurrence of urban traits in a territory characterized by a high degree of industrialization which may have the center of its seat in the city after all. Only if type construction is to looked upon as completely unrelated to empirical science can it be accepted as a challenge at all. Even type construction, however, has now to be relegated to speculation and armchair science if unsupported by empirical proof.

The contribution by Albert C. Reisse, Jr. is mentioned because it is unique and falls somewhere outside the ordinary. There are many other contributions which simply assemble the knowledge now available on different aspects of urbanism. Pigget discusses the "City in Action Civilization" and we have a discussion of the "Adaptation of Design to the Metropolis" by Richard J. Neutra. In neither of these quite adequate contributions do we meet more than the assembly of already well-known materials. The science of urbanism is not very much advanced by these contributions. Still, the summary statements are probably the best written with all this material at hand. They will continue to be used by the student for what they are: competent undertakings to gather all available knowledge in brief summary, and to recapitulate it. The book thus yields the platform from which new contributions may have to take over.

SVEND RIEMER

University of California at Los Angeles

Urban Mortgage Lending—Comparative Markets and Experience. By J. E. Morton. A Study by the National Bureau of Economic Research, New York. Princeton: Princeton University Press, 1956. pp. xxi, 187. $4.00.

This competent summary and comparative analysis is the culminating item in the Urban Real Estate Finance Project series of monographs initiated in 1945. It is Dr. Morton's role to consolidate and compare the findings of the following companion studies, carrying his account of the urban mortgage investment market through 1953: *Urban Mortgage Lending by Life Insurance Companies*, by R. J. Saulnier, *Commercial Bank Activities in Urban Mortgage Financing*, by Carl F. Behrens, *History and Policies of the Home Owners' Loan Corporation*, By C. Lowell Harriss, and *Urban Real Estate Markets: Characteristics and Financing*, by Ernest M. Fisher. In addition, the author drew upon the unpublished work of Professor E. E. Edwards on savings and loan associations, John Lintner's study of mutual savings banks, a special sample survey of the portfolios of institutional lenders as of 1946-47, and the 1950 Census Survey of Residential Financing. The product of those labors is a skillfully drawn picture of the urban mortgage investment market as a whole which highlights significant characteristics and trends.

The first three chapters are descriptive of the supply side of the mortgage market—the level and changes in nonfarm mortgage debt, the structure of the market in terms of institutional distribution, regional distribution and distribution by type for property mortgage.

Chapter 4 describes the characteristics of outstanding mortgages with respect to type of property, government insurance status, contract length, loan-to-value ratios, interest rates and loan age distribution. Marked changes over time in these attributes are revealed.

The last chapter deals with mortgage lending experience. Mortgage investors will find the analysis of the determinants of foreclosure and the treatment of loss experience to be particularly illuminating and useful.

Dr. Morton's contribution in this summary monograph lies not only in its technical excellence but also in its selective and penetrating analysis. To him and to the other authors in the real estate finance series goes the gratitude of both academicians and business men for this comprehensive treatment of a subject field so long neglected. The National Bureau of Economic Research is to be commended for its sponsorship of these greatly needed studies.

RICHARD U. RATCLIFF
University of Wisconsin

Index to

LAND ECONOMICS

a quarterly Journal of

PLANNING, HOUSING & PUBLIC UTILITIES

VOLUME XXXII

February 1956-November 1956

LEADING ARTICLES

REPORTS AND COMMENTS

BOOK REVIEWS

SUBJECT MATTER

Among articles to appear in February 1957 issue:

Land Tenure in the Hashemite Kingdom of Jordan....*By Gabriel Baer*

The Role of the Courts in Public Utility Regulation as Exemplified in Oregon....*By Ralph C. Hoeber*

The New Jersey Turnpike....*John J. George* and *Thomas W. Ryley*

L

P

C

D

R

C

T

P

M

F

M

"
W
E
W

VOLU

LAND ECONOMICS

a quarterly journal of

PLANNING, HOUSING & PUBLIC UTILITIES

"WHATEVER MAY BE THE LIMITATIONS WHICH TRAMMEL INQUIRY ELSEWHERE WE BELIEVE THE GREAT STATE UNIVERSITY OF WISCONSIN SHOULD EVER ENCOURAGE THAT CONTINUAL AND FEARLESS SIFTING AND WINNOWING BY WHICH ALONE THE TRUTH CAN BE FOUND."

From Report of Board of Regents, University of Wisconsin, 1894

VOLUME XXXII, NUMBER 4 NOVEMBER 1956 PRICE $2.00 A COPY

AMERICAN ECONOMIC REVIEW

Contents

Volume XLVI September 1956

The AMERICAN ECONOMIC REVIEW, a quarterly, is the official publication of the American Economic Association and is sent to all members. The annual dues are $6.00. Address editorial communications to Dr. Bernard F. Haley, Editor, American Economic Review, Stanford University, Room 220, Stanford, California; for information concerning other publications and activities of the Association, communicate with the Secretary-Treasurer, Dr. James Washington Bell, American Economic Association, Northwestern University, Evanston, Illinois. Send for information booklet.